

412TW-PA-16435



Sample Size Effects on Bias in CE50 and CE90

TODD G. REMUND

**AIR FORCE TEST CENTER
EDWARDS AFB, CA**

SEPTEMBER 2016

**4
1
2
T
W**

**Approved for public release; distribution is unlimited.
412TW-PA-16435**

**412TH TEST WING
EDWARDS AIR FORCE BASE, CALIFORNIA
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE**

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) 09-09-2016		2. REPORT TYPE White Paper		3. DATES COVERED (From - To)	
4. TITLE AND SUBTITLE Sample Size Effects on Bias in CE50 and CE90				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Todd G. Remund				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) AND ADDRESS(ES) 812 TSS/ENT 307 E. Popson Ave Edwards AFB, CA 93524				8. PERFORMING ORGANIZATION REPORT NUMBER 412TW-PA-16435	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S) N/A	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release A: distribution is unlimited.					
13. SUPPLEMENTARY NOTES CA: Air Force Test Center Edwards AFB CA CC: 012100					
14. ABSTRACT Small samples affect estimates of the ninetieth and fiftieth percentiles for circular error. This report details this bias behavior for various possible target location error distributions for various sample sizes and exhibits nonparametric percentile estimation behavior.					
15. SUBJECT TERMS Quantile, percentile, bias, sample size, nonparametric, target location error, CEP, CE90, TLE.					
16. SECURITY CLASSIFICATION OF: Unclassified			17. LIMITATION OF ABSTRACT None	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON 412 TENG/EN (Tech Pubs)
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER (include area code) 661-277-8615

Report: Sample Size Effects on Bias in CE50 and CE90

Author: Todd Remund, Edwards AFB, 812 TSS/ENT, Statistical Methods Group

This report focuses on choosing a sample size based on low bias in an estimate of a percentile. As sample size gets larger the bias in estimating CE50 and CE90 gets smaller, note a nonparametric estimator is used for estimating the percentile of circular errors.

Table 1: Bivariate Normal Parameter Combinations

Run	Standard Deviation		Mean		rho
	X	Y	X	Y	
1	20	20	0	0	0
2	10	20	0	0	0
3	20	10	0	0	0
4	10	10	0	0	0
5	20	20	20	0	0
6	10	20	20	0	0
7	20	10	20	0	0
8	10	10	20	0	0
9	20	20	0	20	0
10	10	20	0	20	0
11	20	10	0	20	0
12	10	10	0	20	0
13	20	20	20	20	0
14	10	20	20	20	0
15	20	10	20	20	0
16	10	10	20	20	0
17	20	20	0	0	0.7
18	10	20	0	0	0.7
19	20	10	0	0	0.7
20	10	10	0	0	0.7
21	20	20	20	0	0.7
22	10	20	20	0	0.7
23	20	10	20	0	0.7
24	10	10	20	0	0.7
25	20	20	0	20	0.7
26	10	20	0	20	0.7
27	20	10	0	20	0.7
28	10	10	0	20	0.7
29	20	20	20	20	0.7
30	10	20	20	20	0.7
31	20	10	20	20	0.7
32	10	10	20	20	0.7

Table 1 contains the description of 32 scenarios describing possible behavior of target location error for targeting pod performance. This is a simulation study of the effects of sample size on bias in percentile estimation where the simulations are based on the parameters in Table 1.

With increasing sample size there will be a point of diminishing returns in reducing the bias of the estimator. An evaluation will be done graphically. The graph of the distributional pattern of the hypothetical bivariate TLE distribution will be given. This is done to indicate that a wide span of scenarios has been considered that should cover basic TLE distributions that may have been encountered in real data.

The study is made up of the parameter combinations for a bivariate normal distribution found in Table 1 above. Bias here is considered to be

$$bias(\widehat{CE}_p) = E(\widehat{CE}_p) - CE_{p,true}.$$

Where the hat symbol represents the estimator and the subscript p denotes a generic percentile.

CE50

In this section a reference line is drawn at a sample of size 12. It may be noted that a sample size smaller than this may be appropriate. The following plots correspond to each of the runs given in table 1.

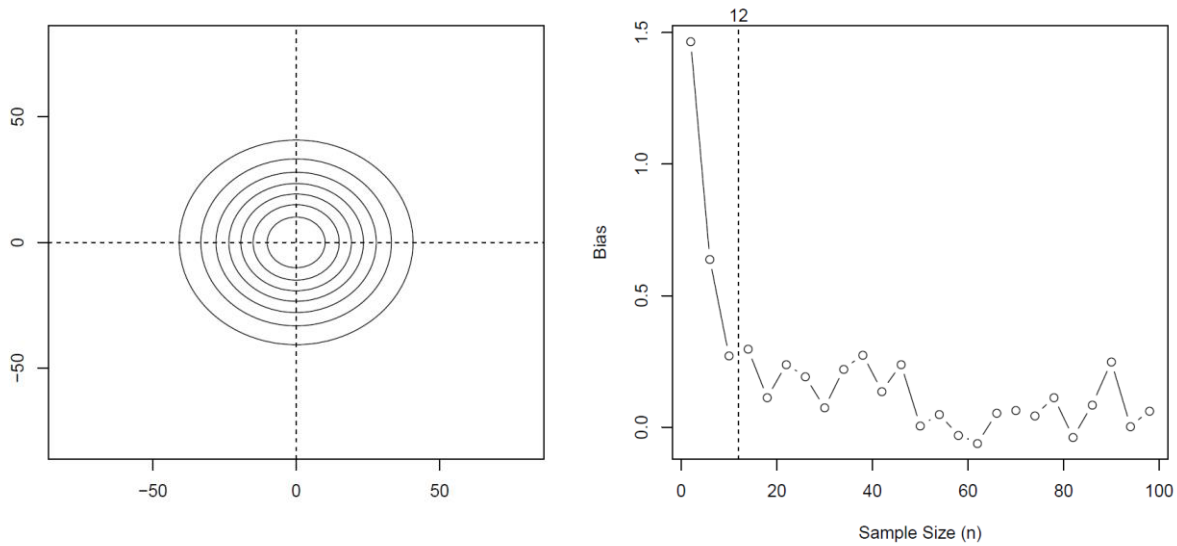


Figure 1 Simulation 1 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

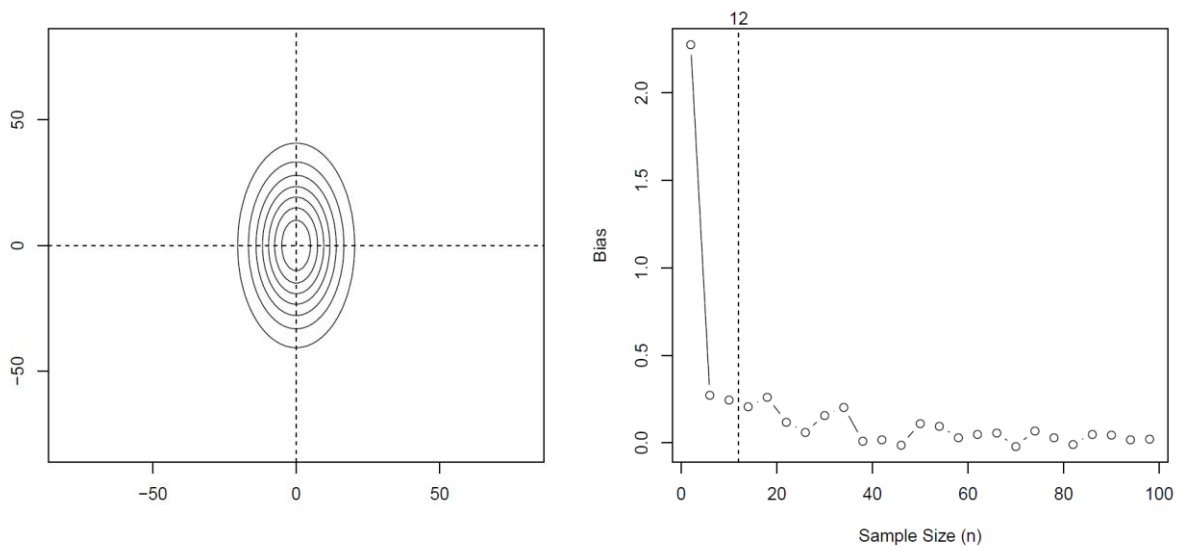


Figure 2 Simulation 2 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

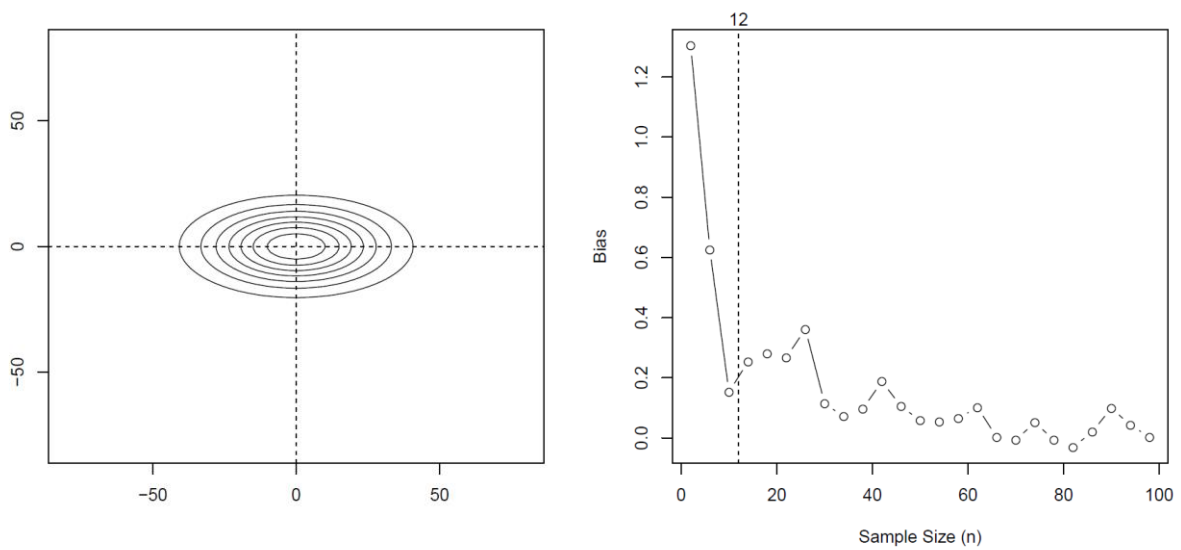


Figure 3 Simulation 3 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

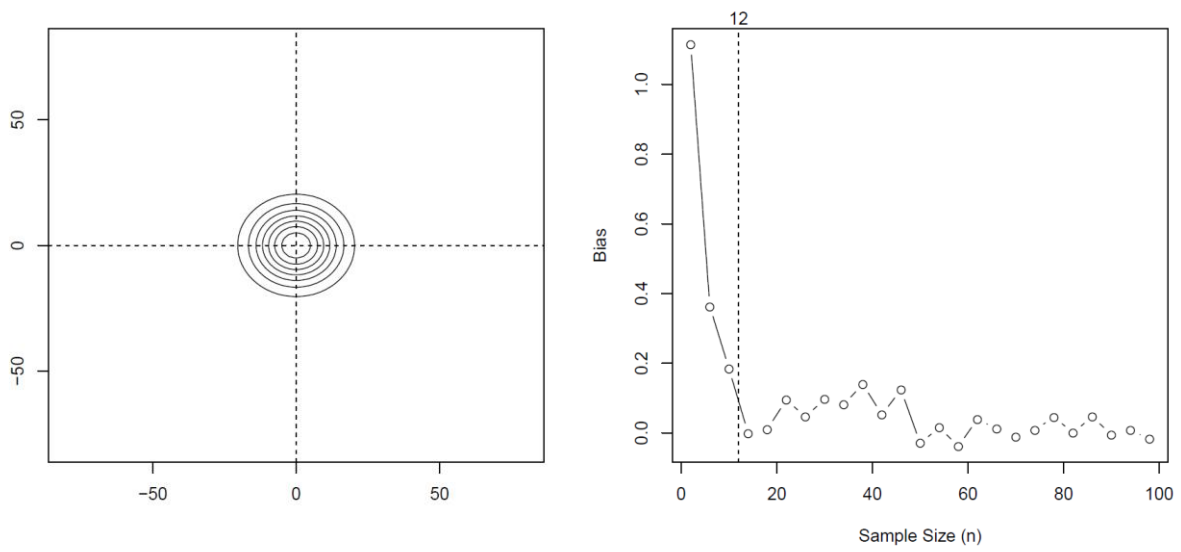


Figure 4 Simulation 4 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

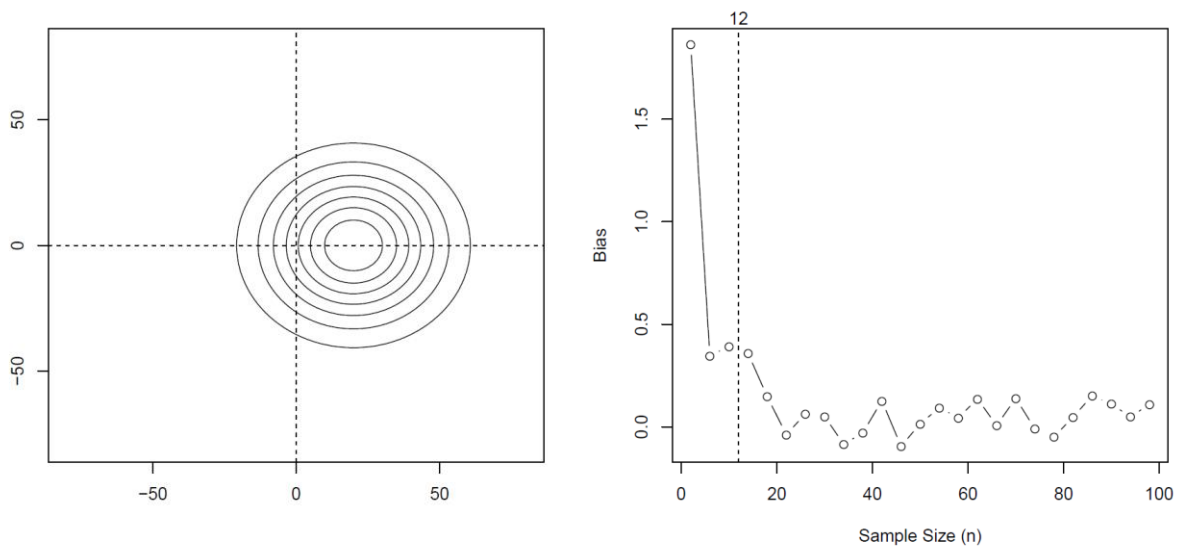


Figure 5 Simulation 5 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

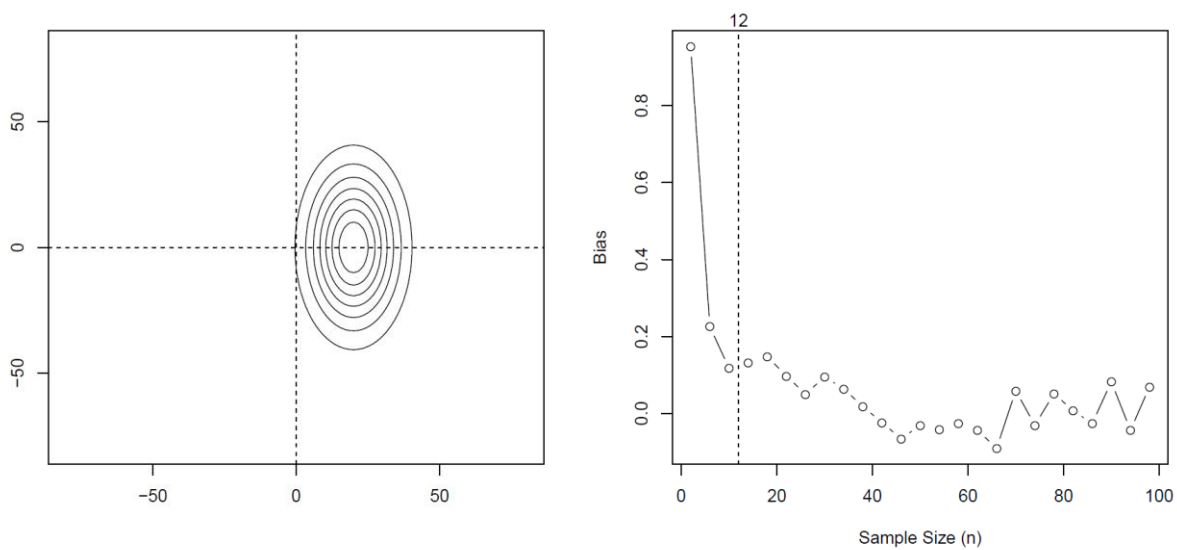


Figure 6 Simulation 6 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

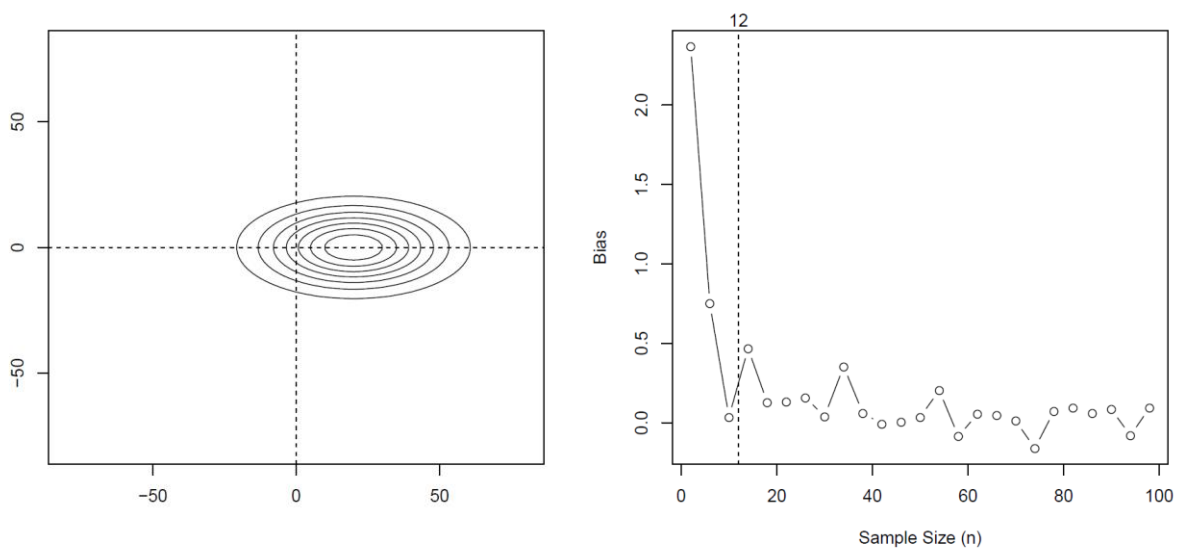


Figure 7 Simulation 7 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

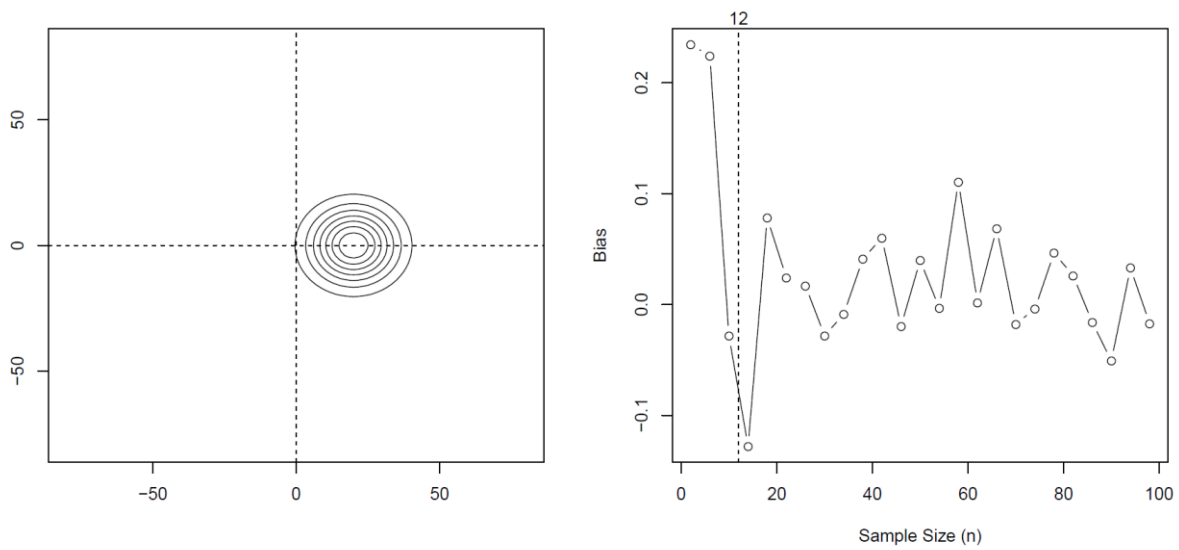


Figure 8 Simulation 8 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

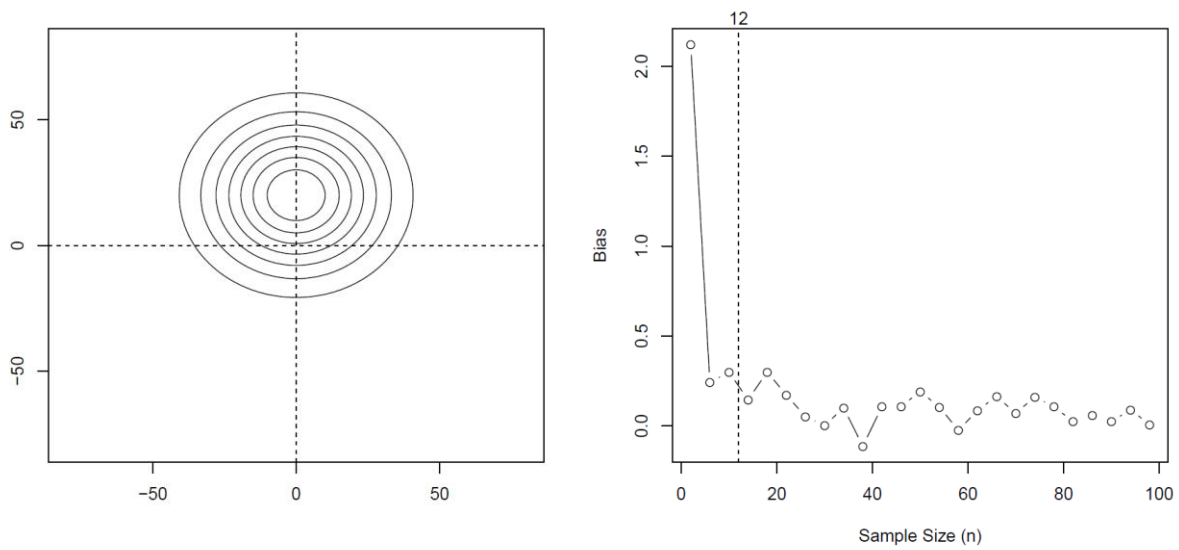


Figure 9 Simulation 9 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

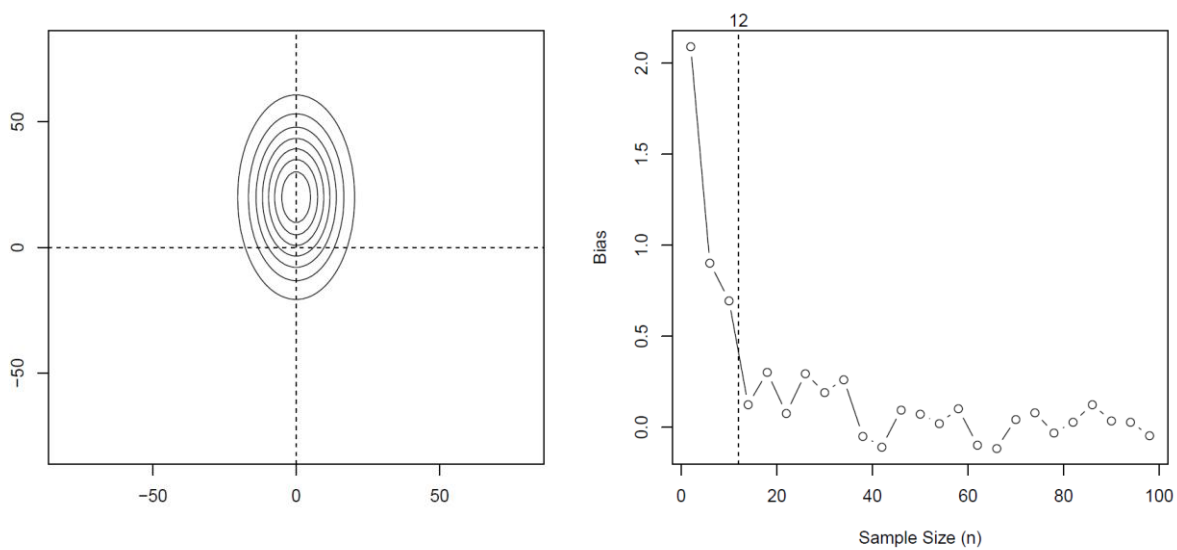


Figure 10 Simulation 10 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

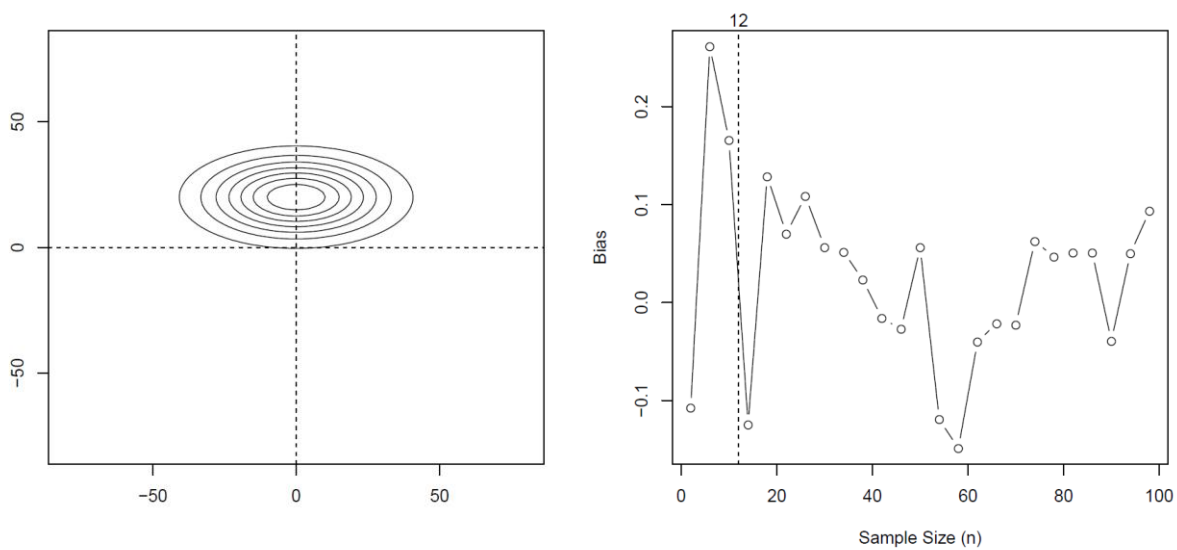


Figure 11 Simulation 11 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

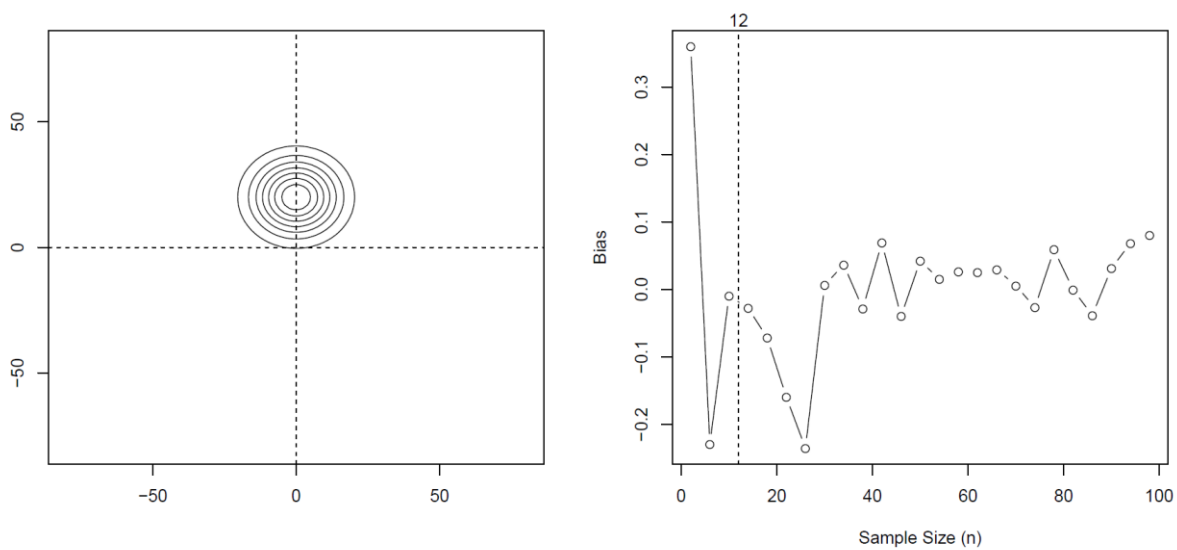


Figure 12 Simulation 12 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

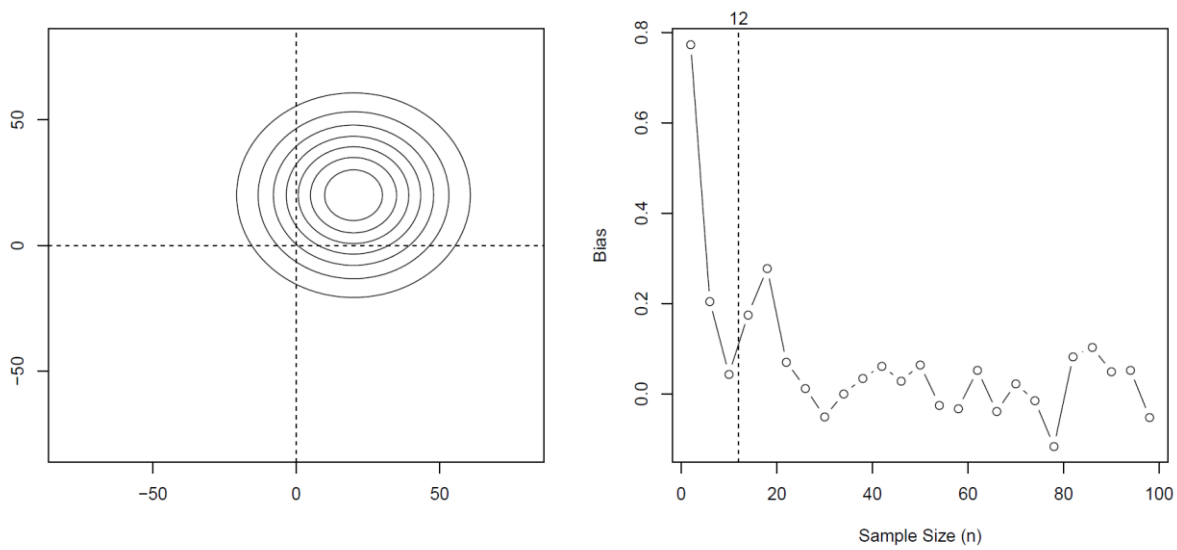


Figure 13 Simulation 13 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

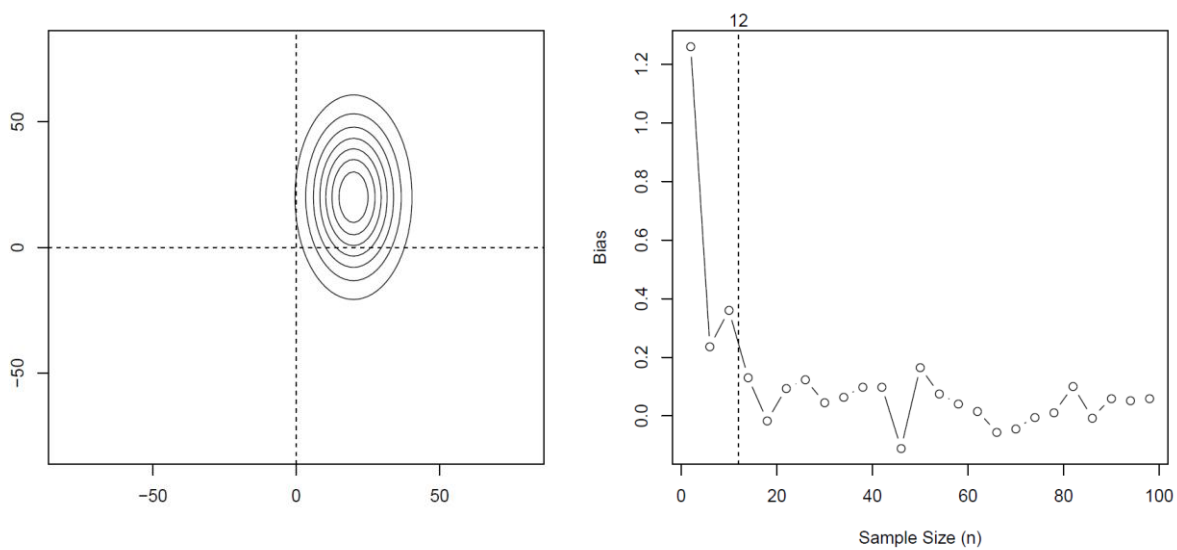


Figure 14 Simulation 14 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

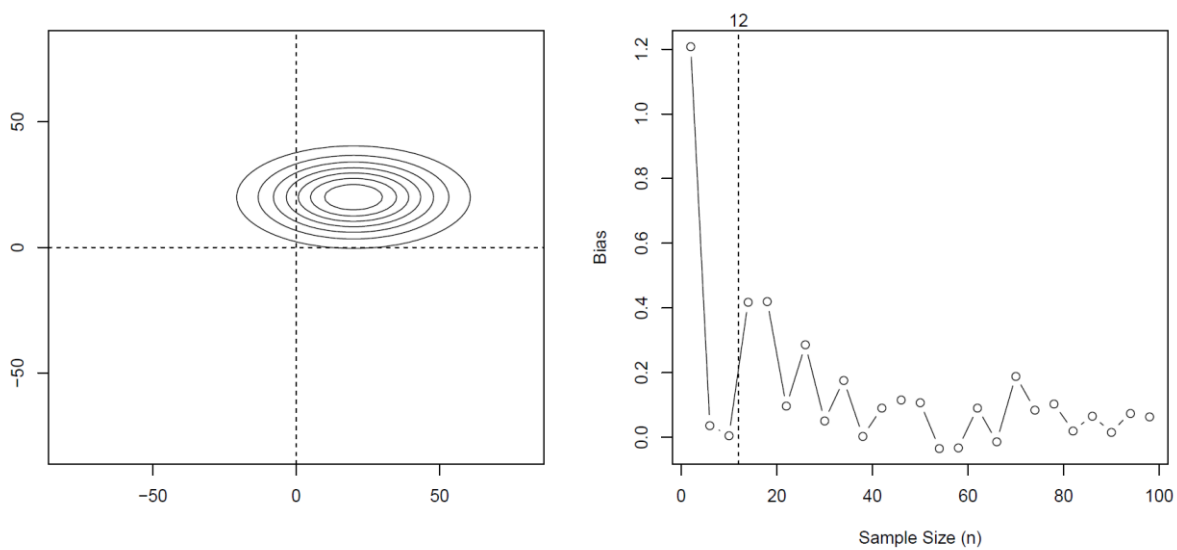


Figure 15 Simulation 15 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

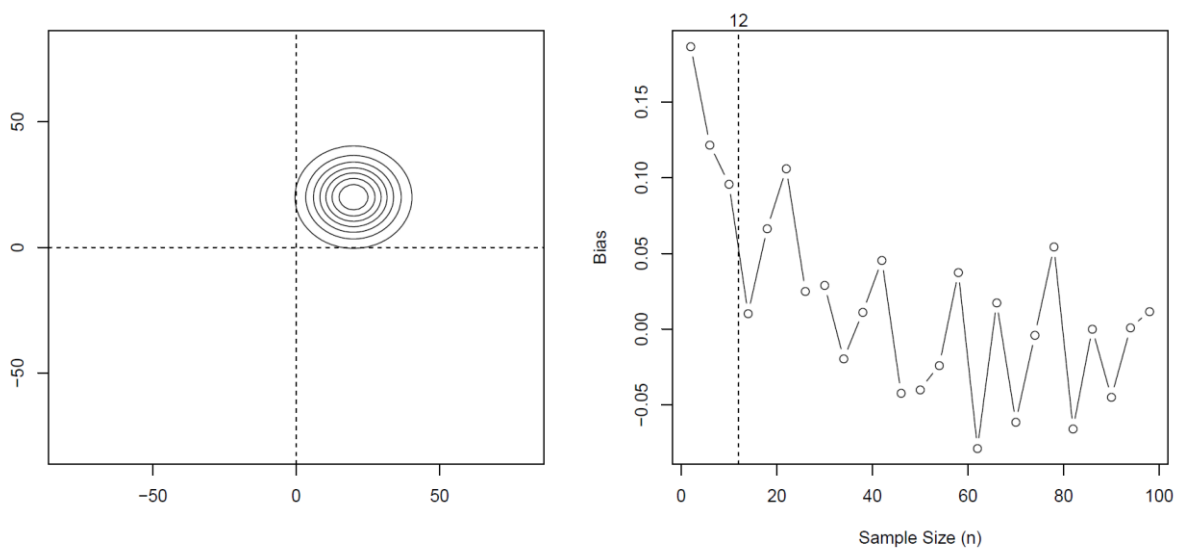


Figure 16 Simulation 16 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

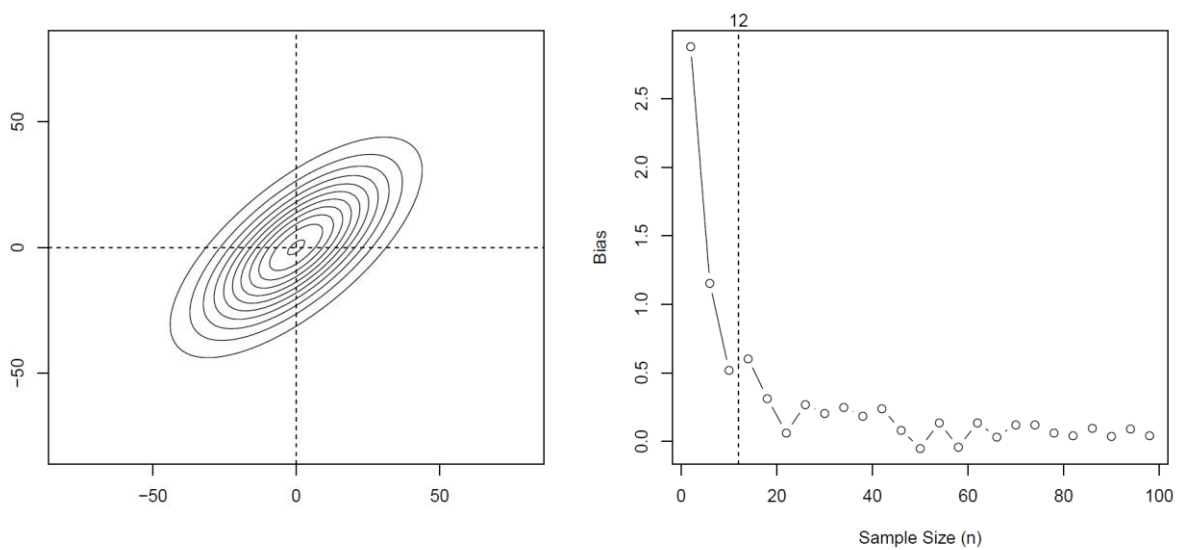


Figure 17 Simulation 17 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

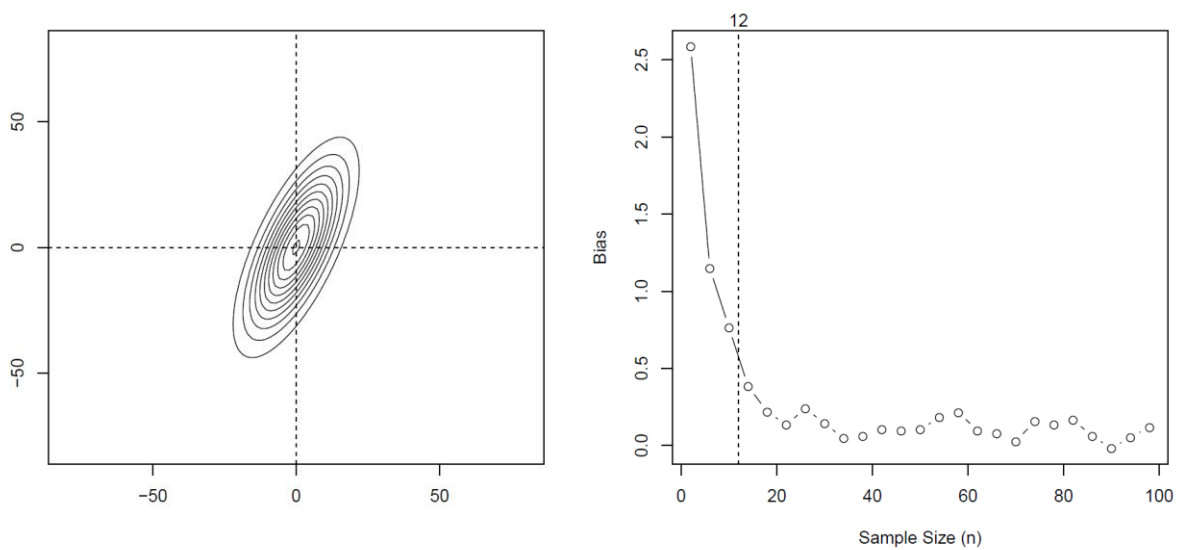


Figure 18 Simulation 18 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

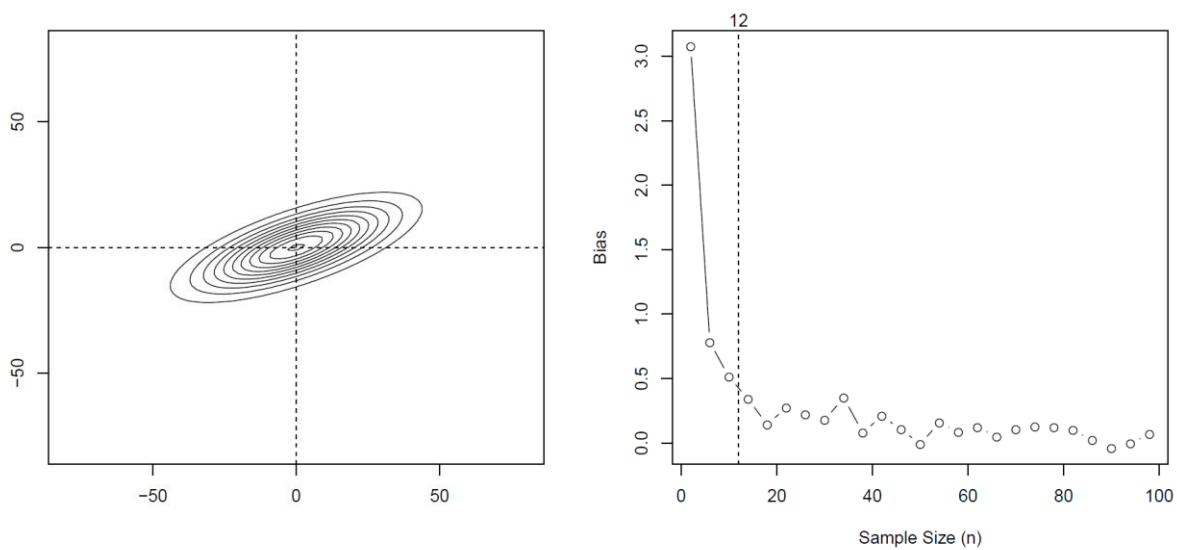


Figure 19 Simulation 19 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

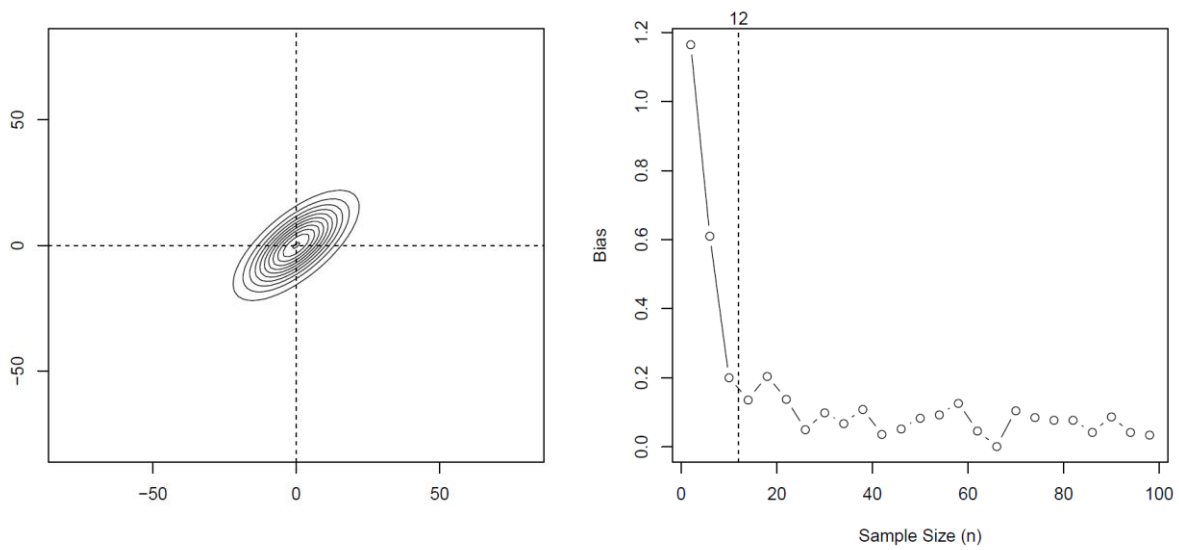


Figure 20 Simulation 20 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

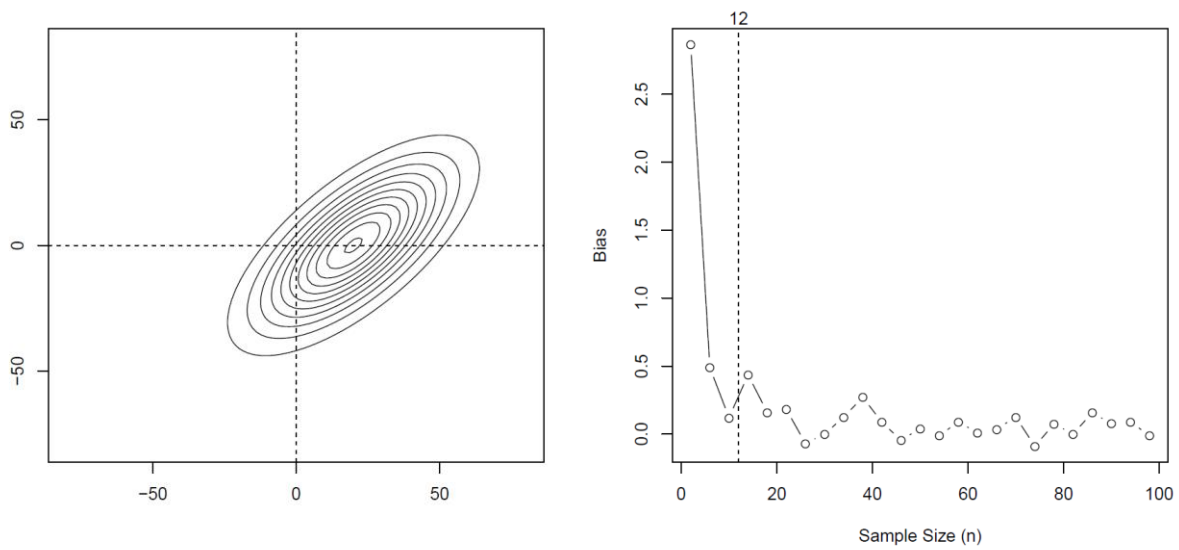


Figure 21 Simulation 21 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

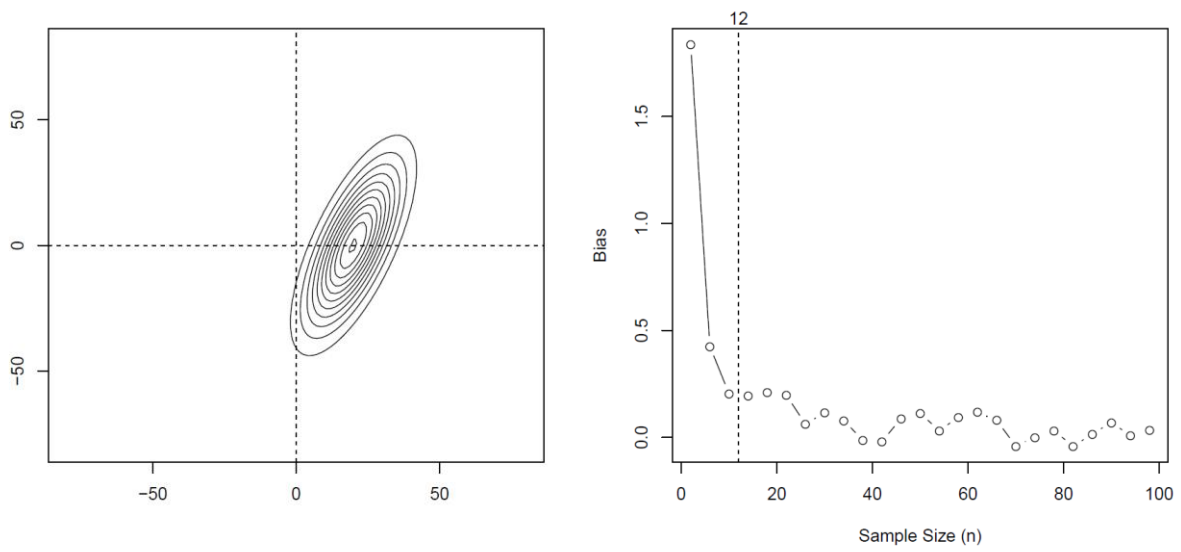


Figure 22 Simulation 22 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

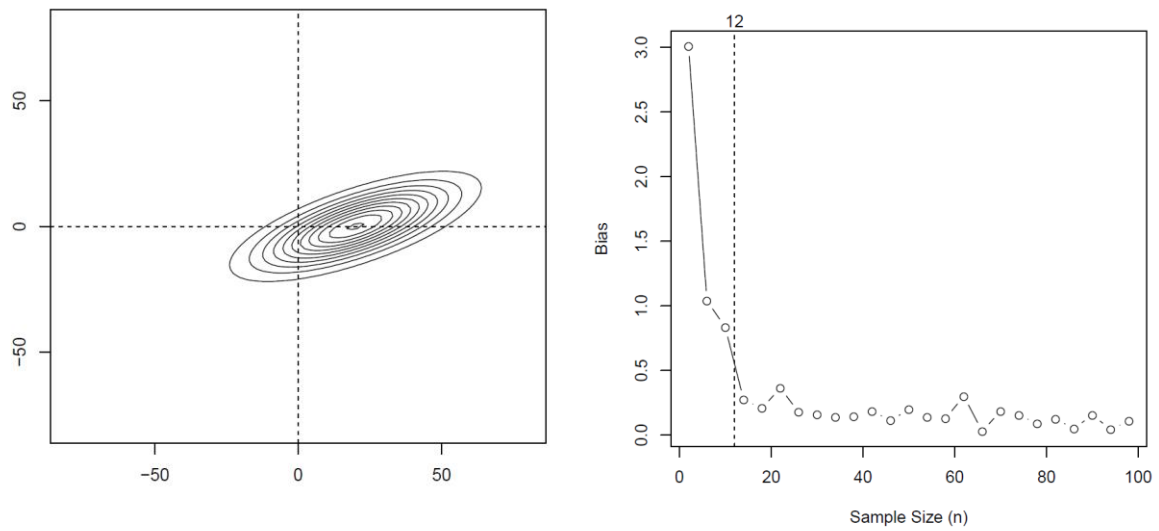


Figure 23 Simulation 23 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

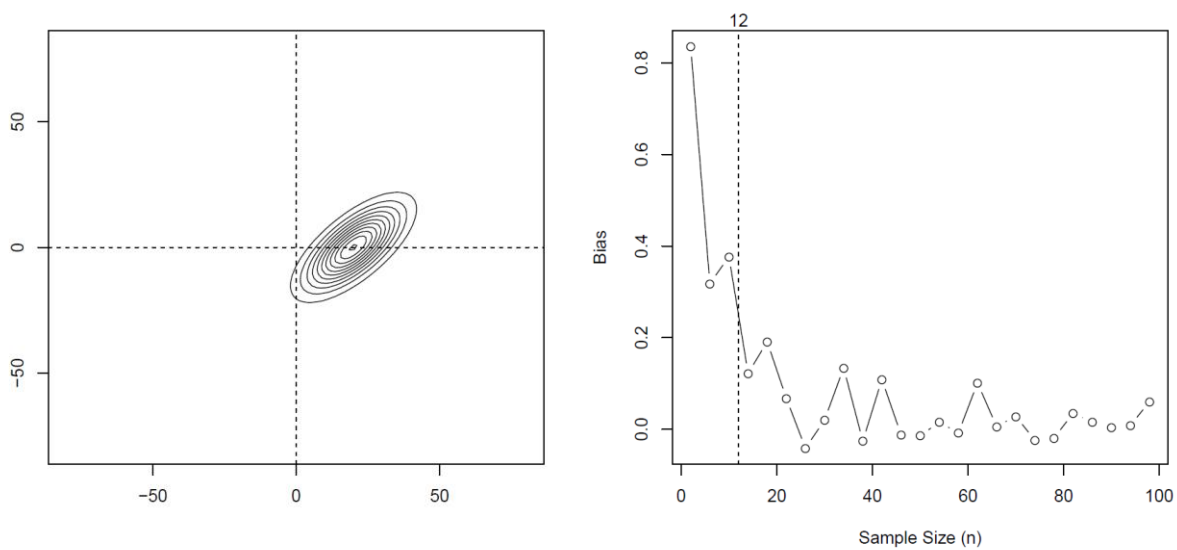


Figure 24 Simulation 24 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

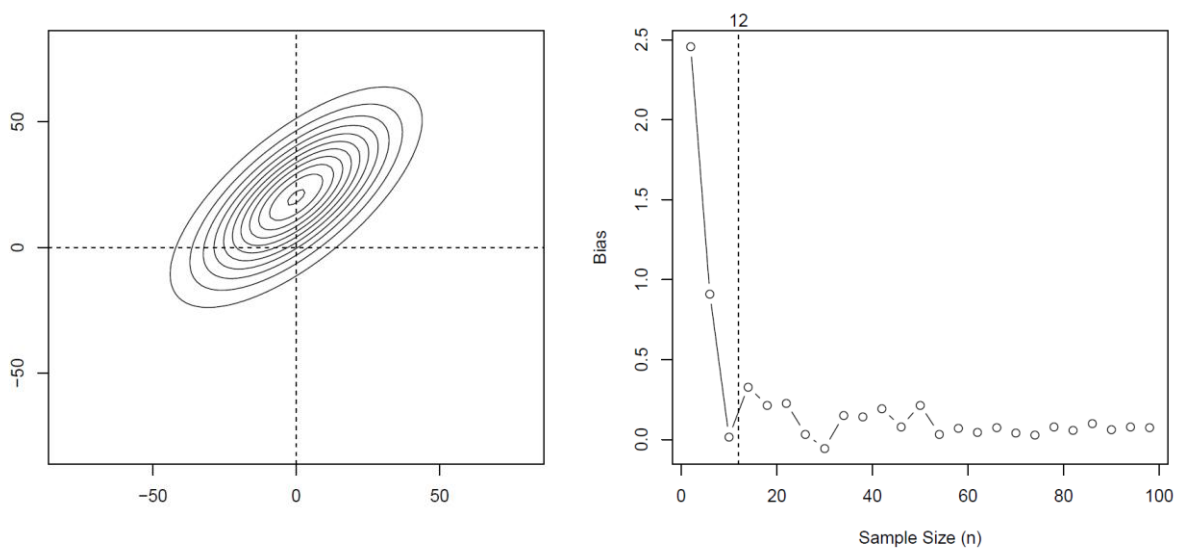


Figure 25 Simulation 25 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

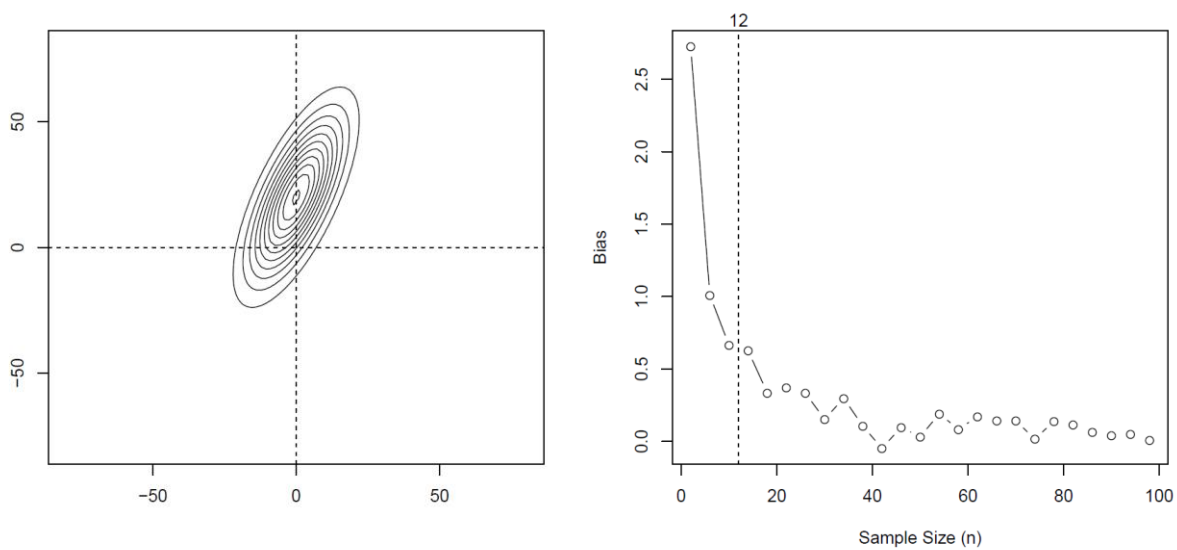


Figure 26 Simulation 26 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

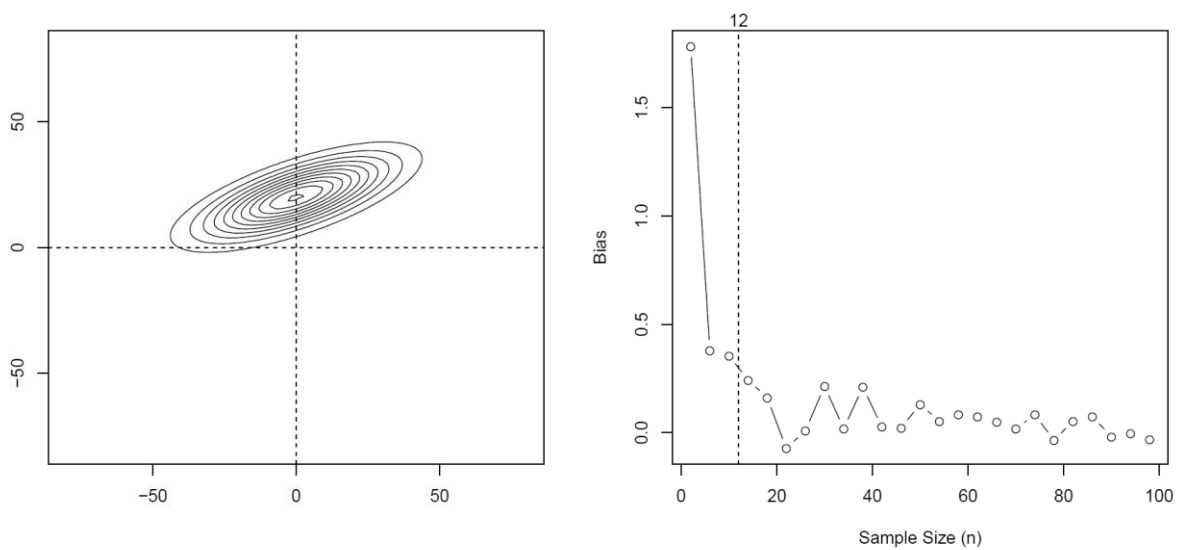


Figure 27 Simulation 27 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

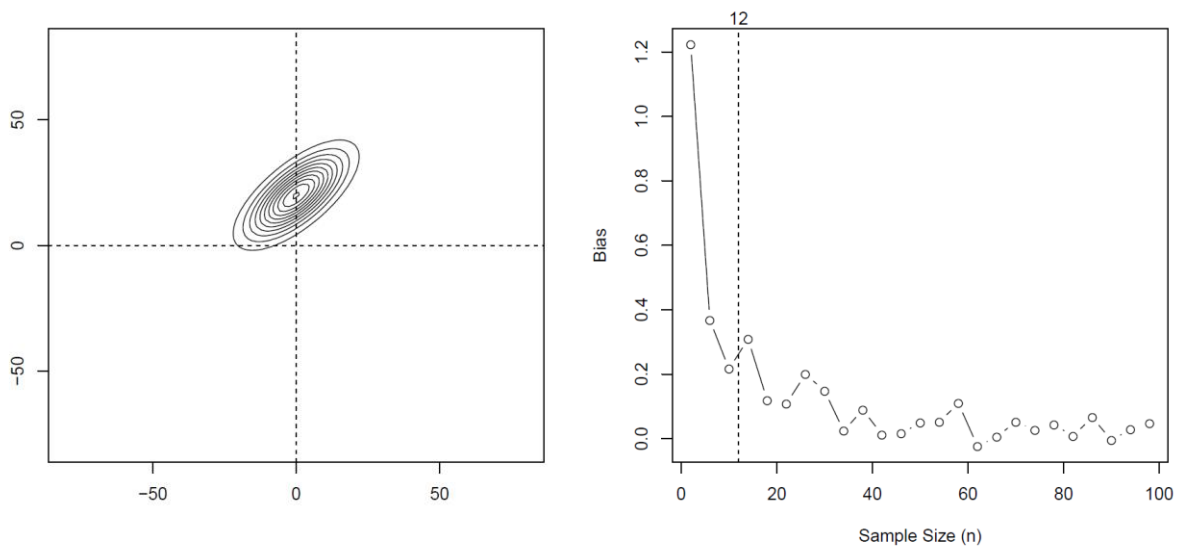


Figure 28 Simulation 28 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

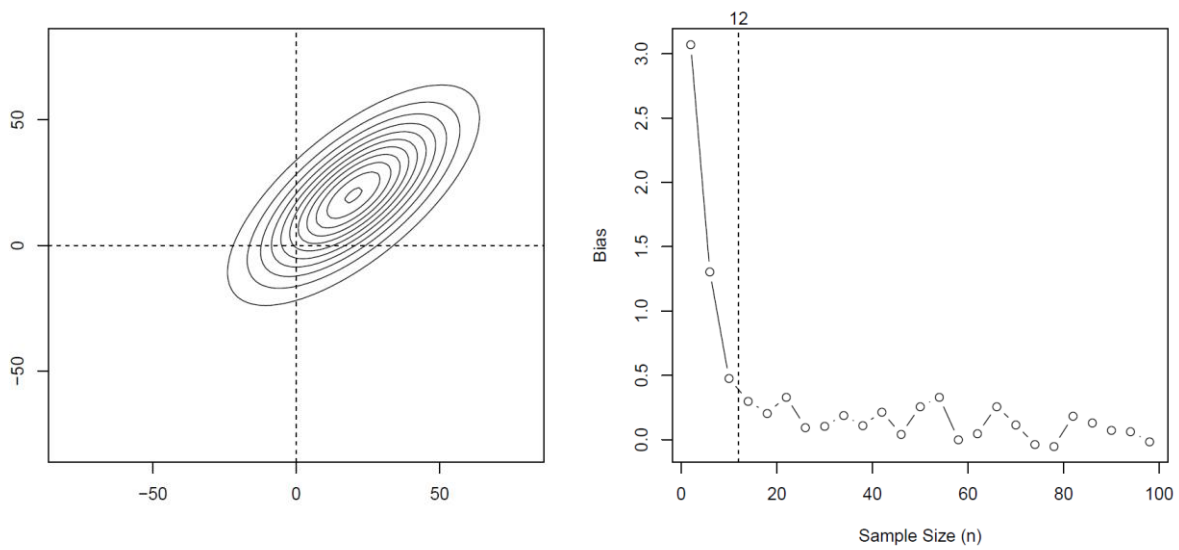


Figure 29 Simulation 29 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

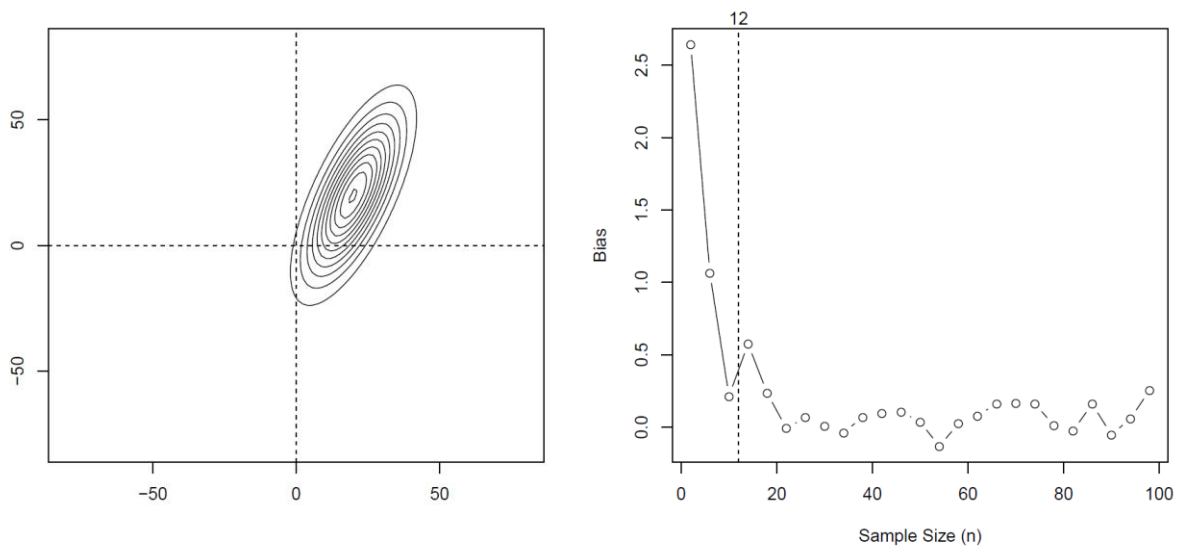


Figure 30 Simulation 30 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

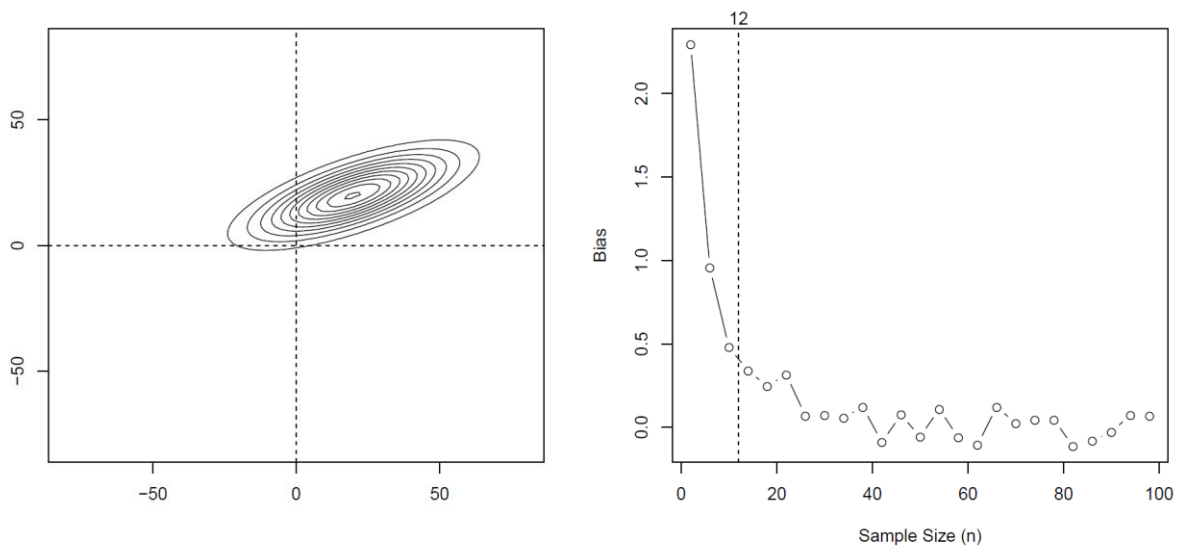


Figure 31 Simulation 31 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

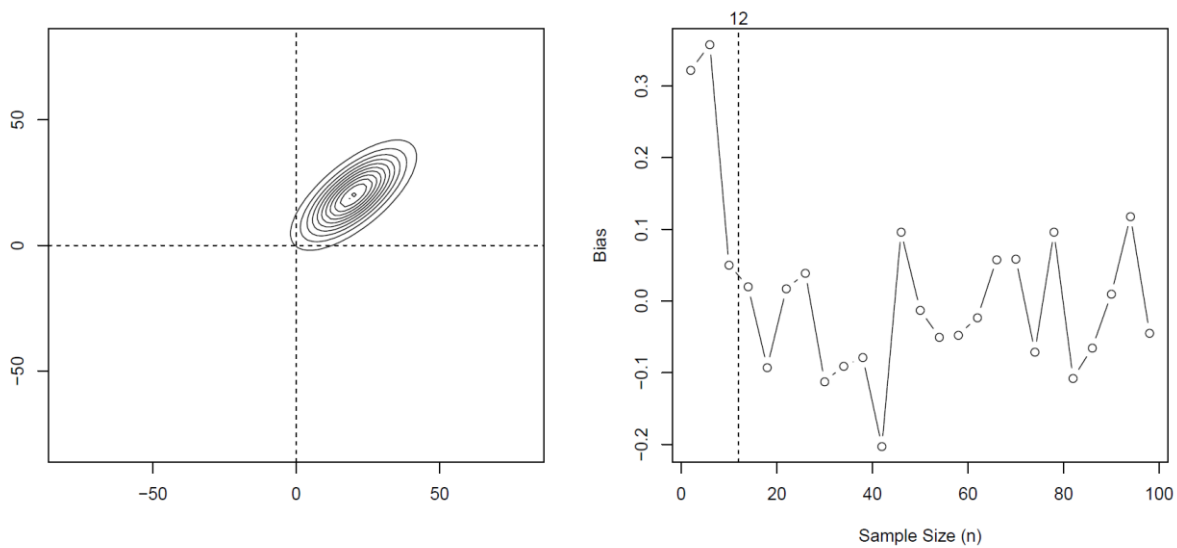


Figure 32 Simulation 32 visual TLE behavior (left), CE50 bias behavior vs. sample size (right).

The bias graphs given for each simulation study indicate that as a general rule, it may be best to go with a sample size somewhere in the neighborhood of 10 to 14 to get a reasonably unbiased estimate of CE50 for the range of normal distribution parameters given in table 1. Also note that the nonparametric estimate for CE50 has a positive bias which indicates that the estimate is biased high for the most part for smaller sample sizes. The size of the bias, even for the smallest sample size, is very small in comparison to any of the distributions standard deviations. Considering the suggested sample sizes between 10 to 14 it is apparent that the bias is very small. For a sample of size 8 it is small enough, however for some of the scenarios there is some reduction in bias that is to be had from sampling a few more samples.

Following are bias estimates for CE90 at various sample sizes, using the same normal distribution parameters given in Table 1.

CE90

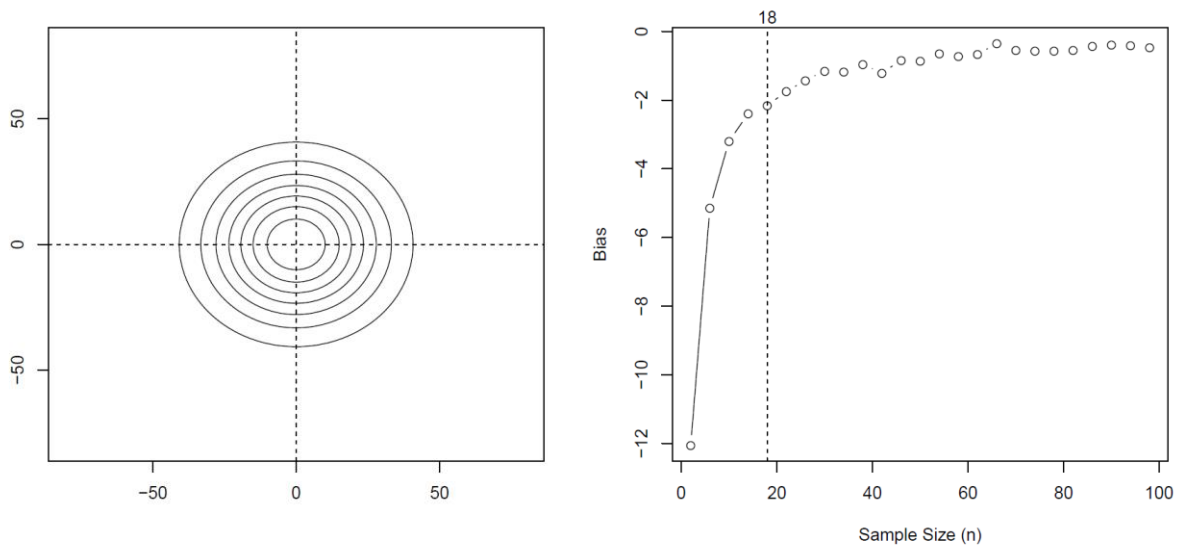


Figure 33 Simulation 1 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

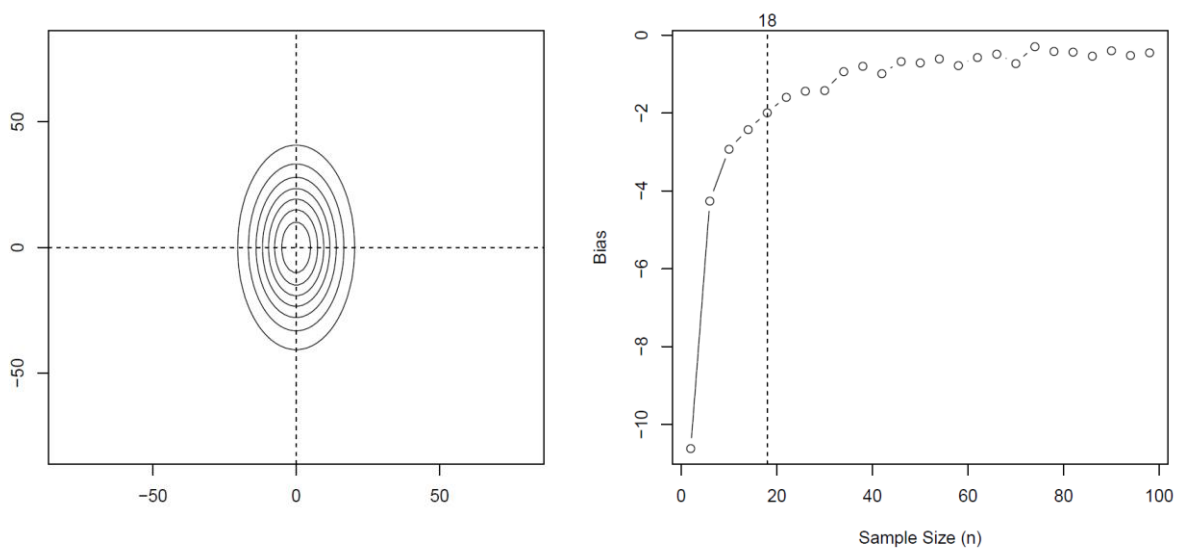


Figure 34 Simulation 2 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

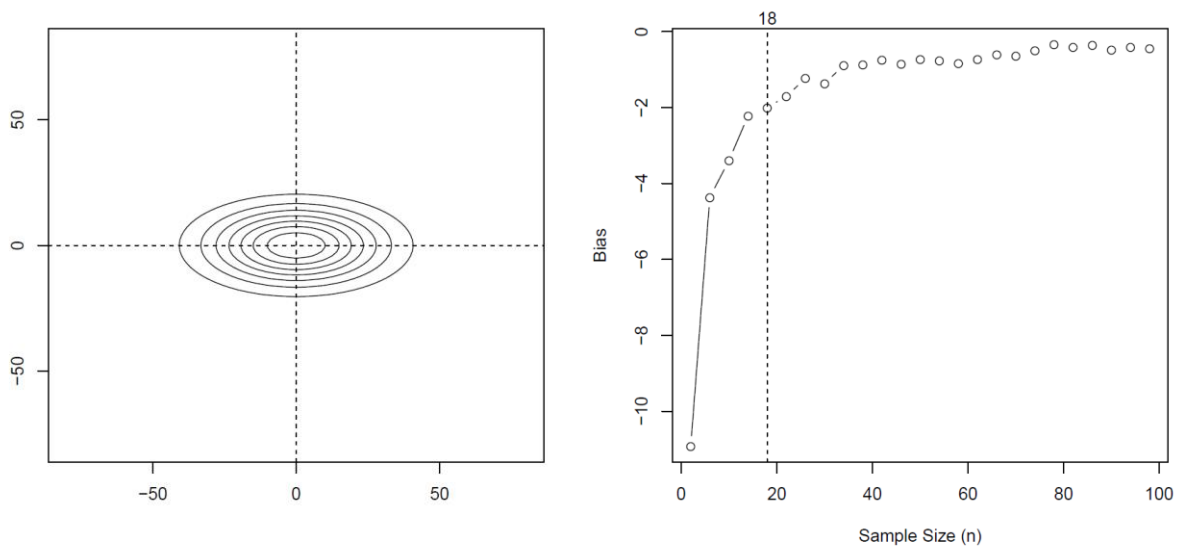


Figure 35 Simulation 3 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

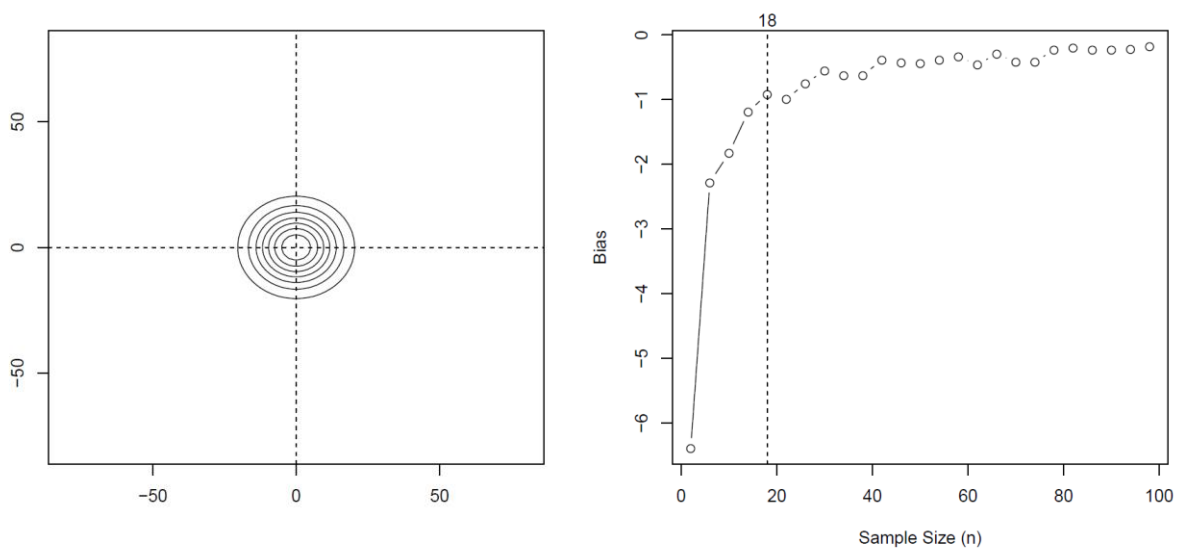


Figure 36 Simulation 4 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

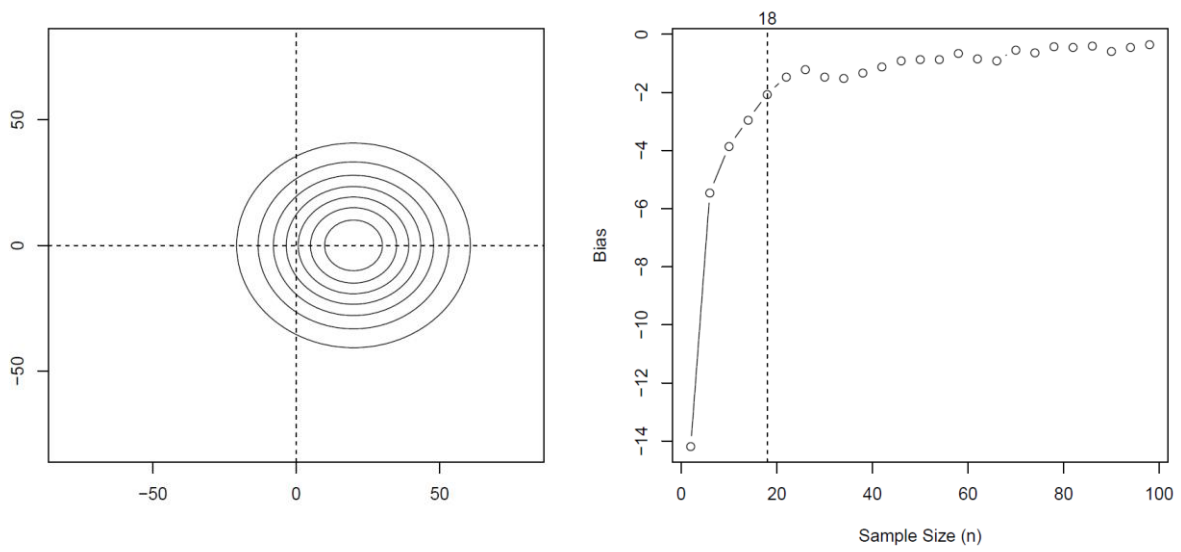


Figure 37 Simulation 5 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

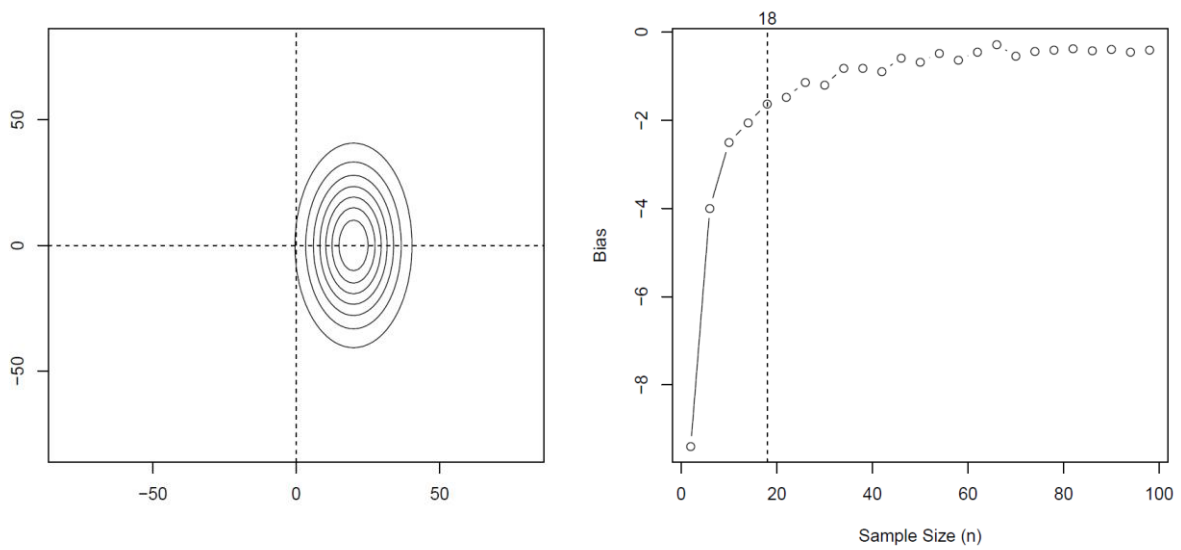


Figure 38 Simulation 6 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

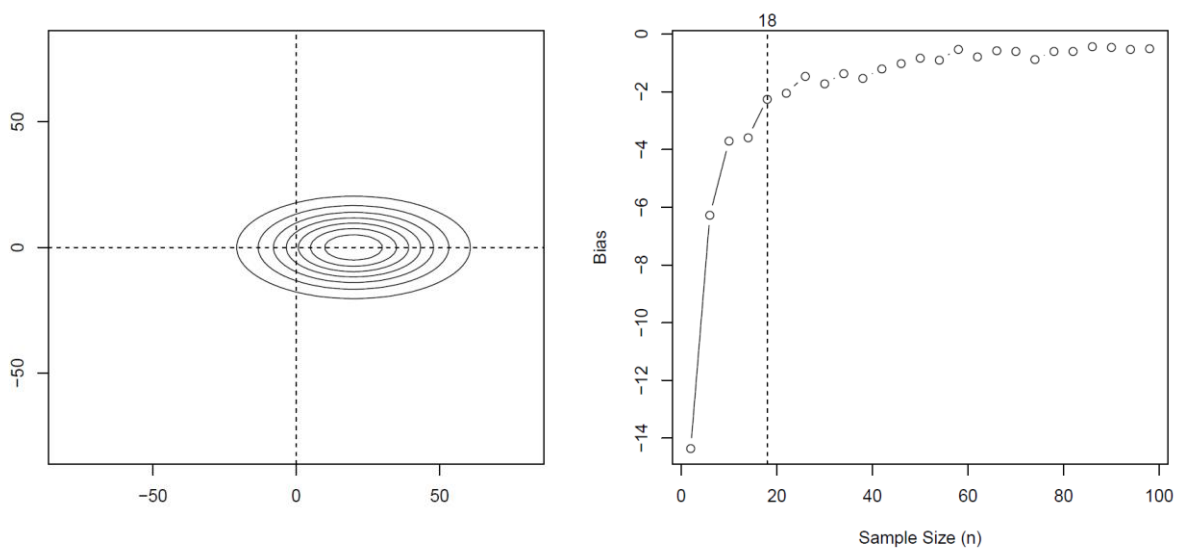


Figure 39 Simulation 7 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

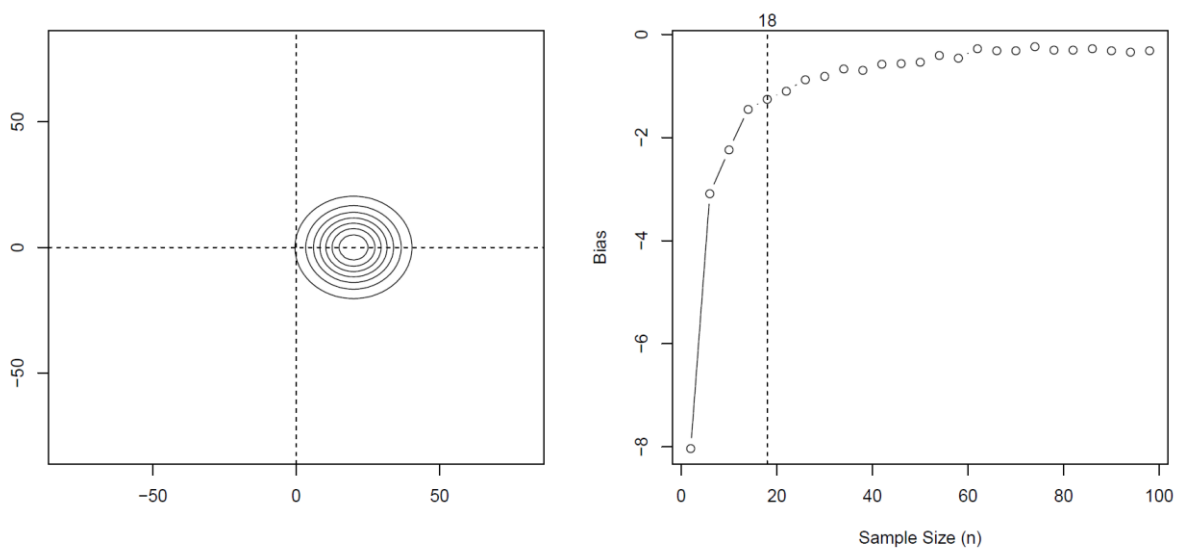


Figure 40 Simulation 8 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

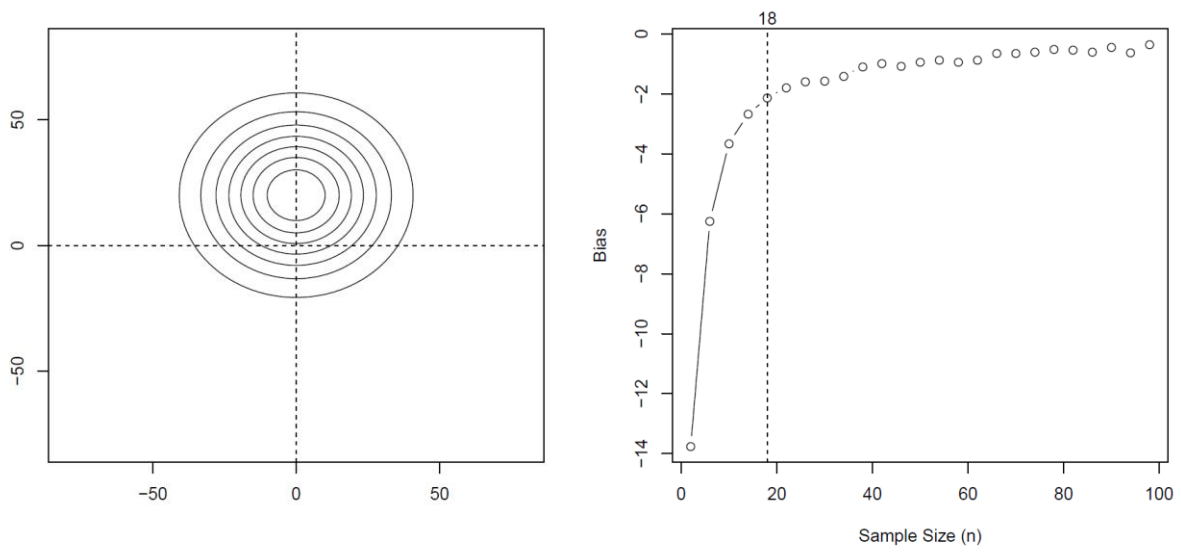


Figure 41 Simulation 9 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

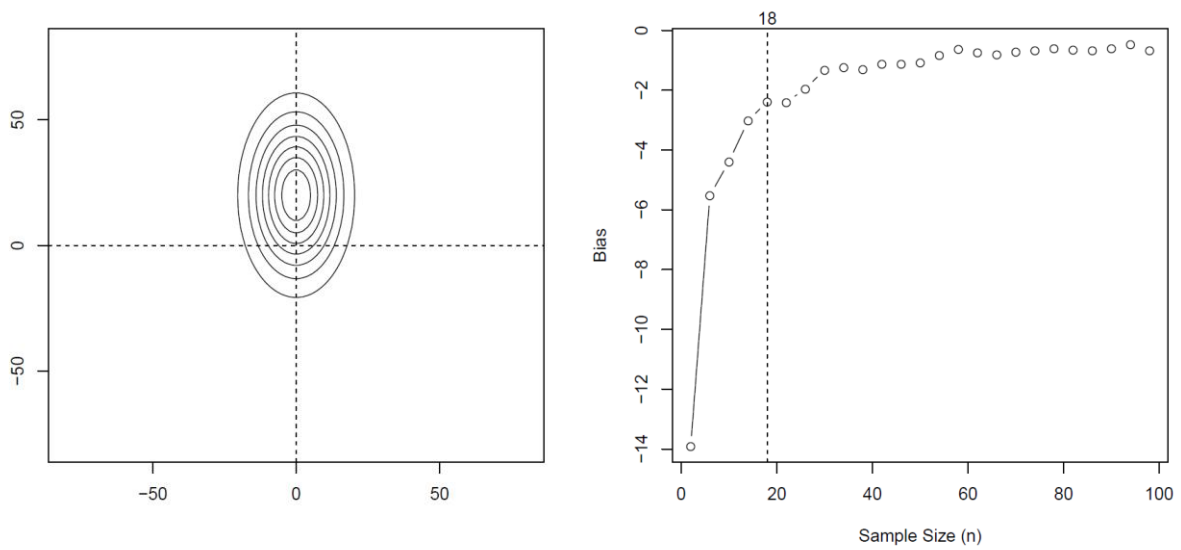


Figure 42 Simulation 10 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

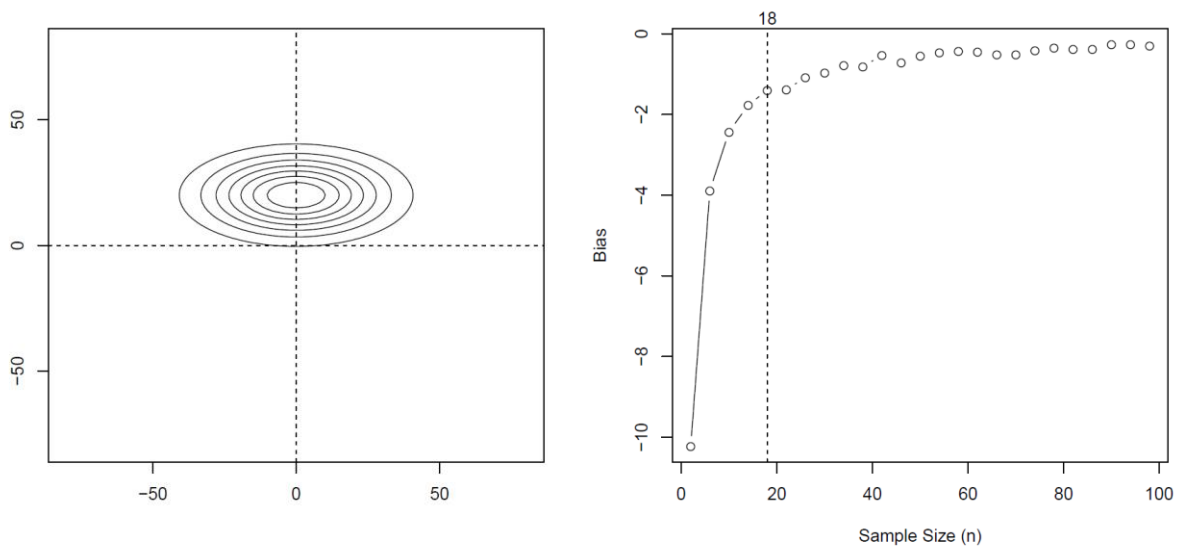


Figure 43 Simulation 11 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

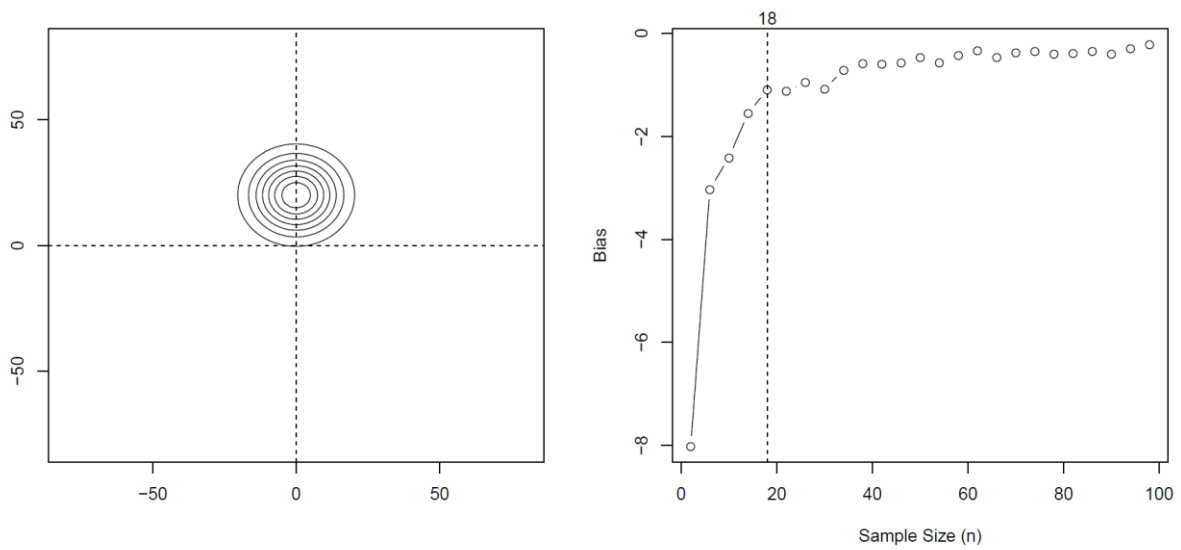


Figure 44 Simulation 12 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

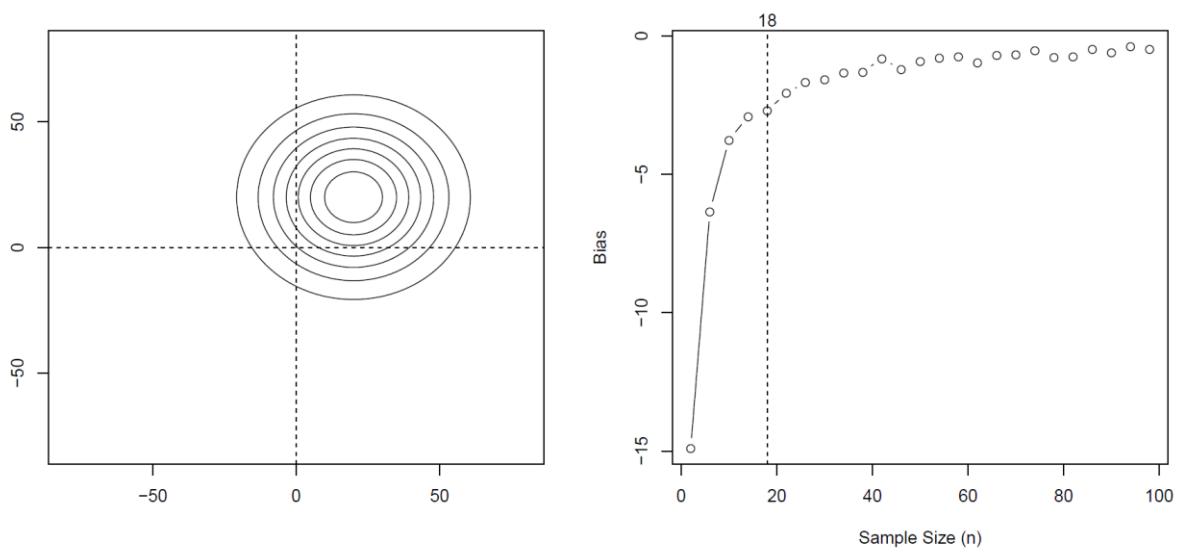


Figure 45 Simulation 13 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

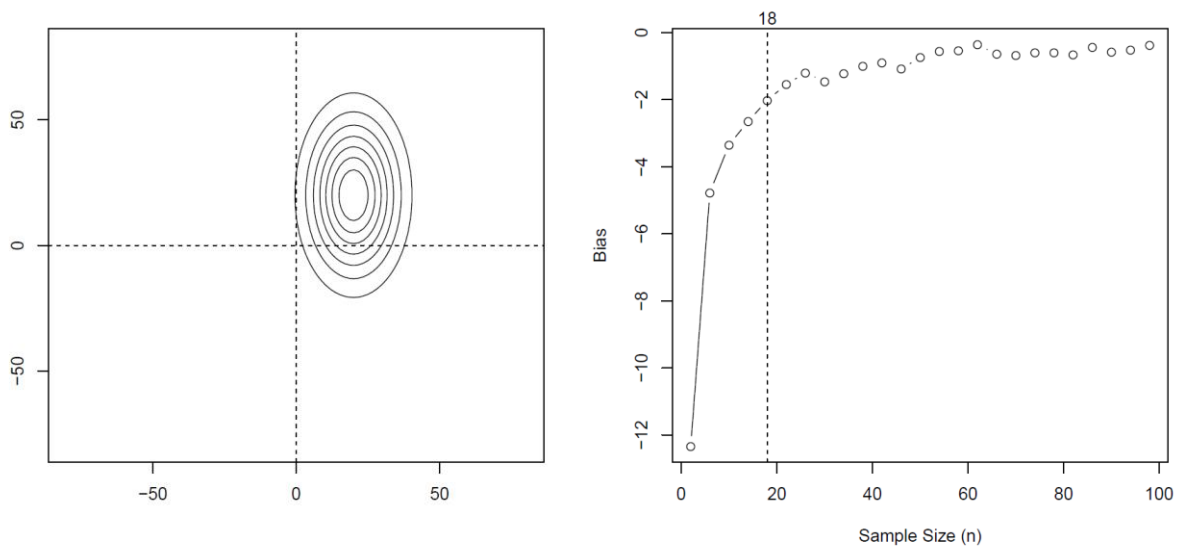


Figure 46 Simulation 14 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

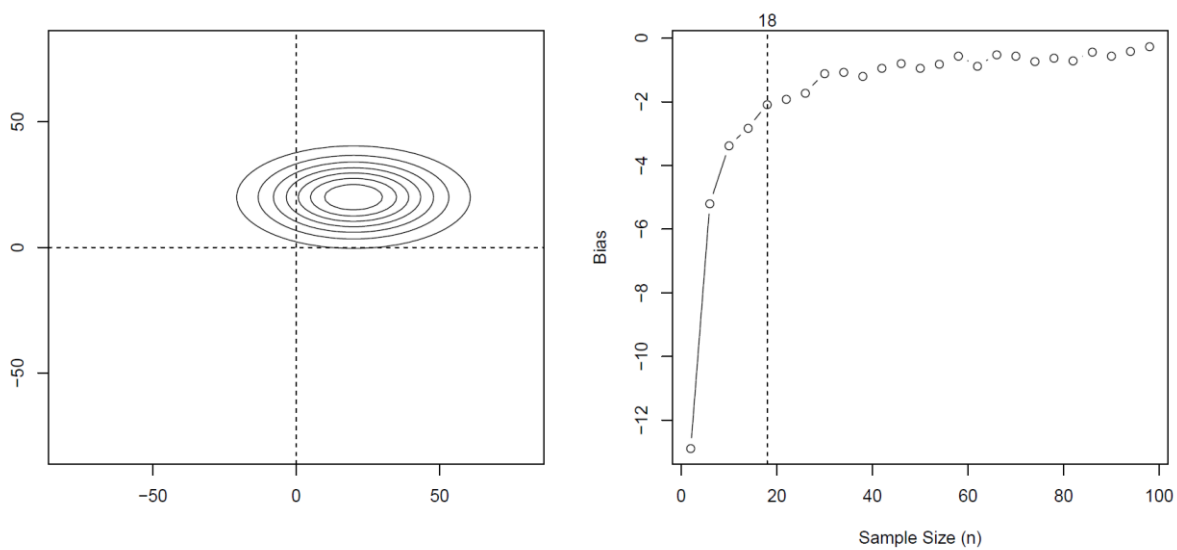


Figure 47 Simulation 15 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

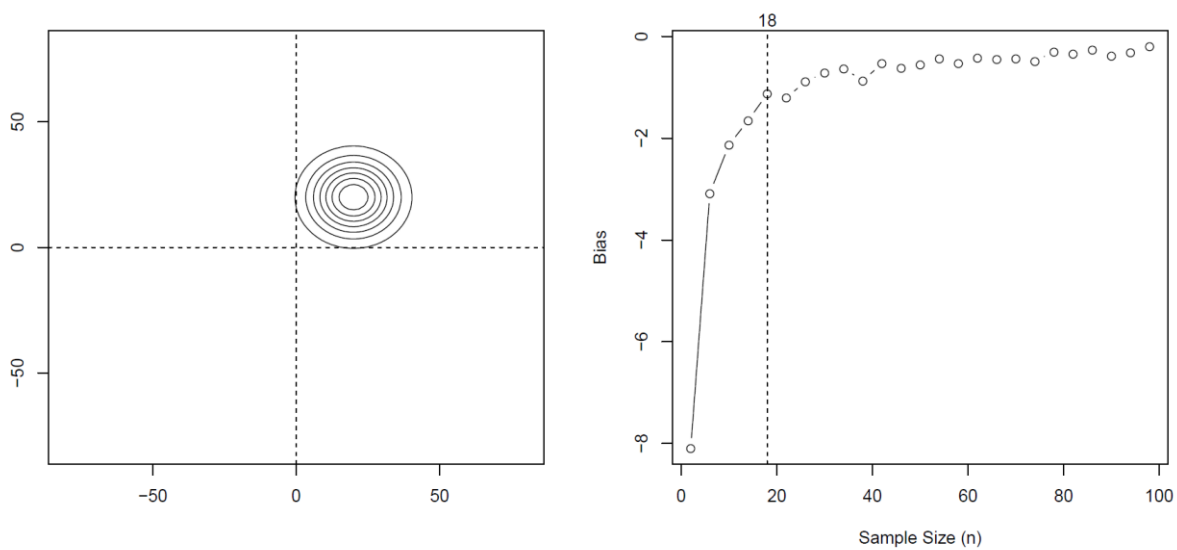


Figure 48 Simulation 16 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

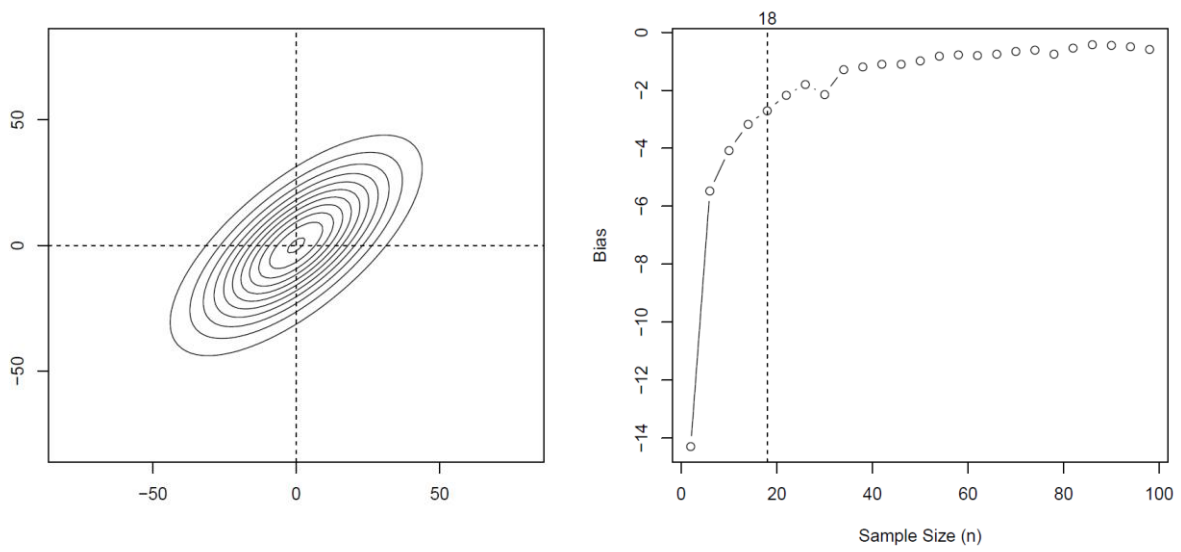


Figure 49 Simulation 17 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

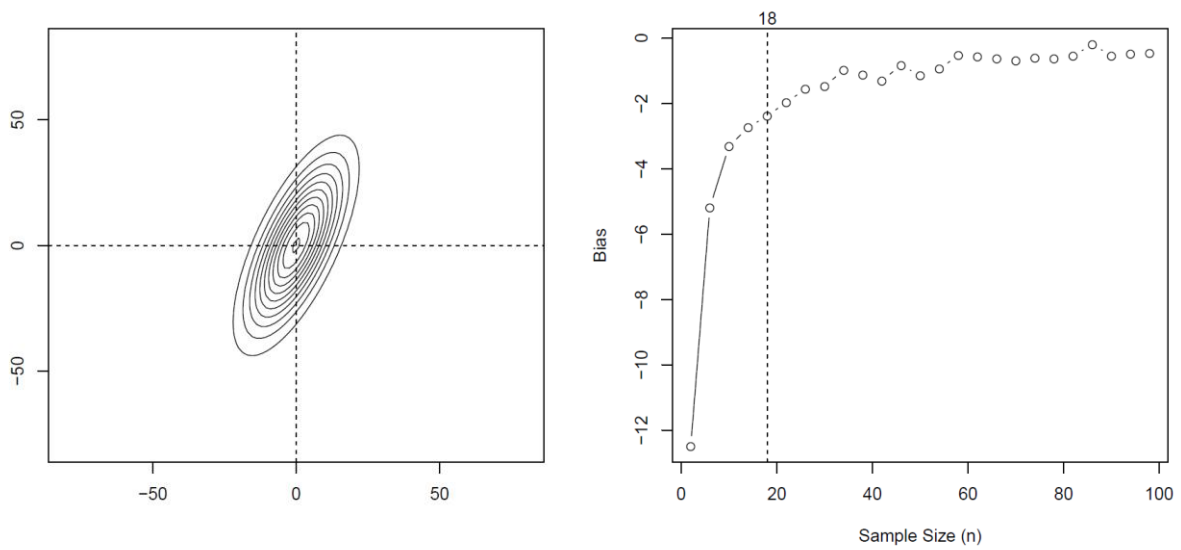


Figure 50 Simulation 18 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

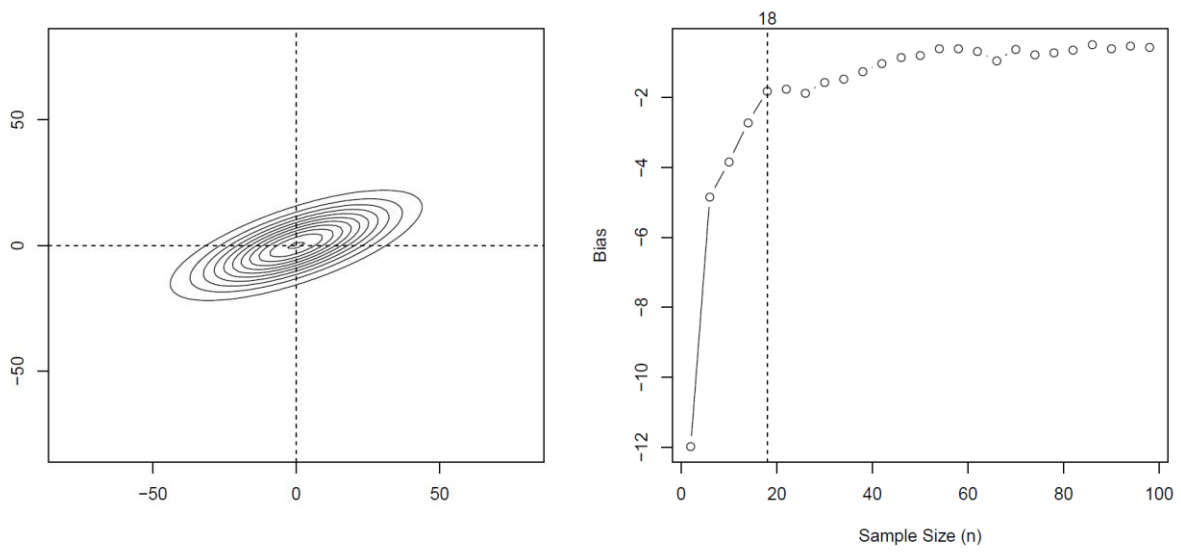


Figure 51 Simulation 19 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

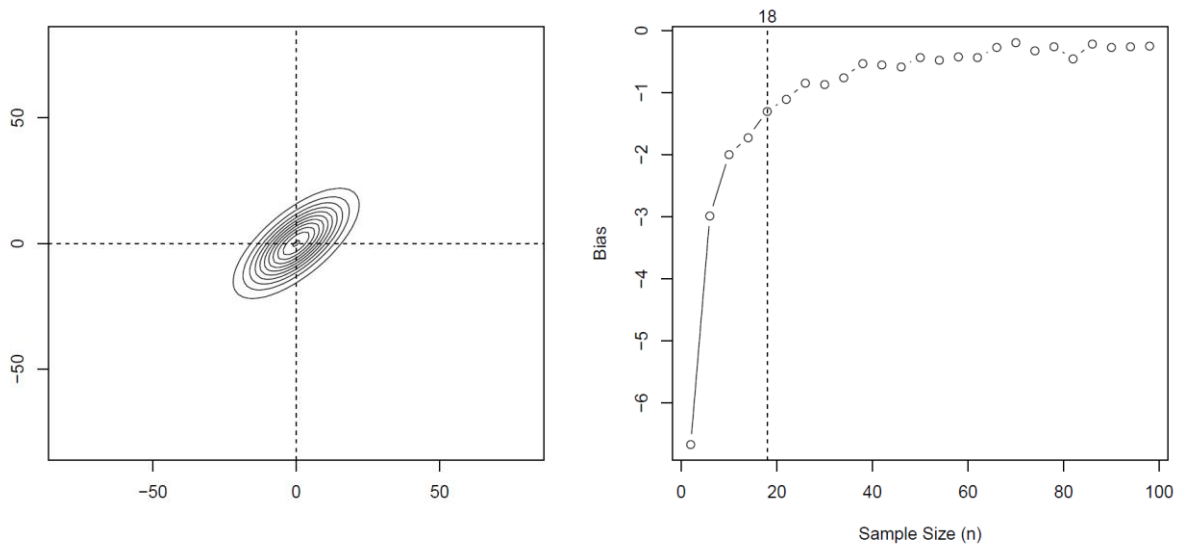


Figure 52 Simulation 20 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

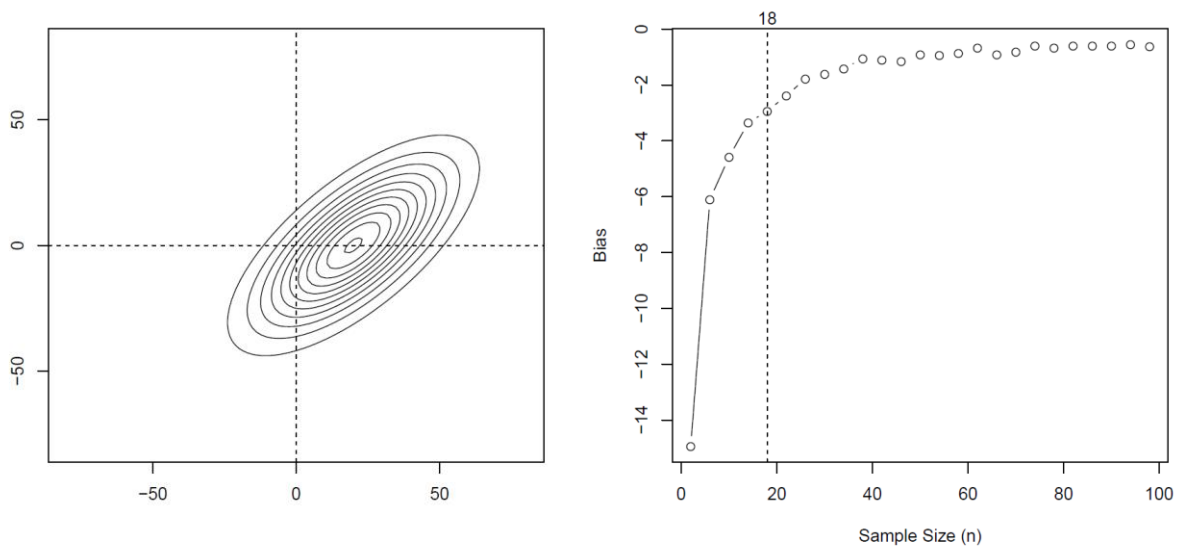


Figure 53 Simulation 21 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

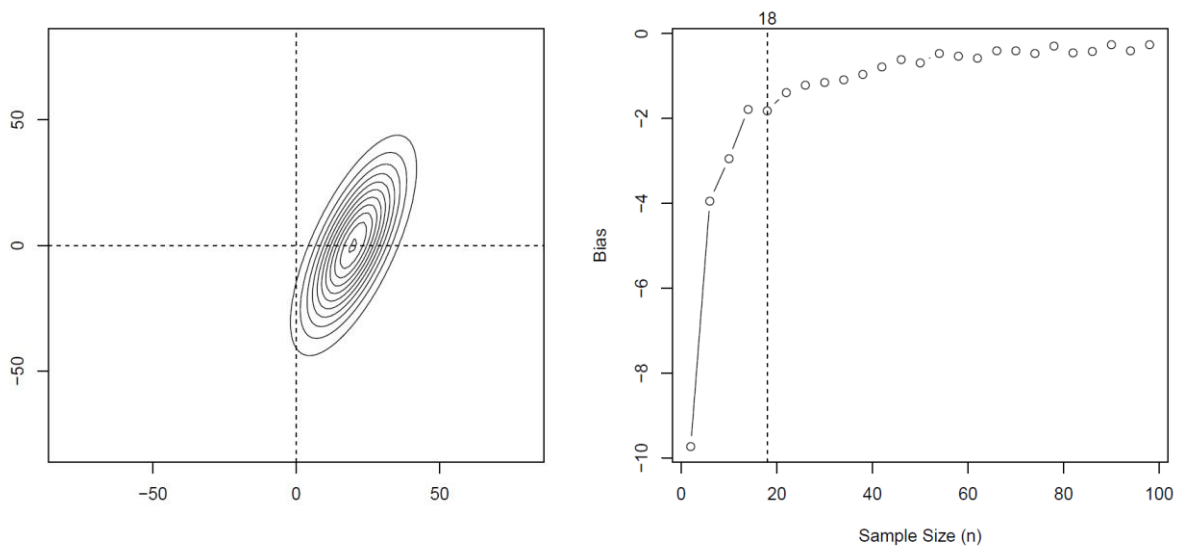


Figure 54 Simulation 22 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

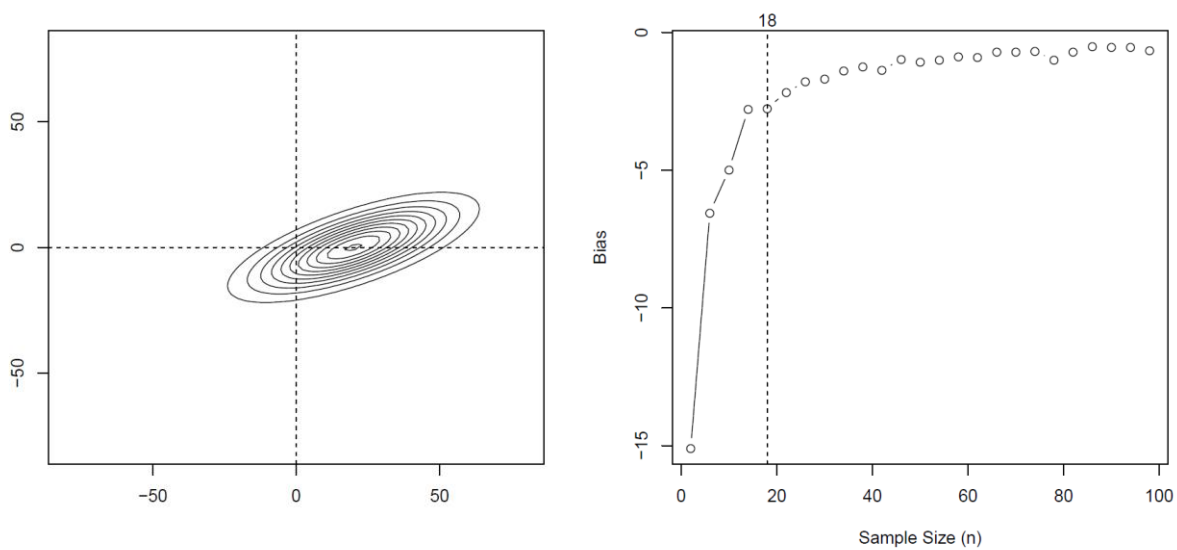


Figure 55 Simulation 23 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

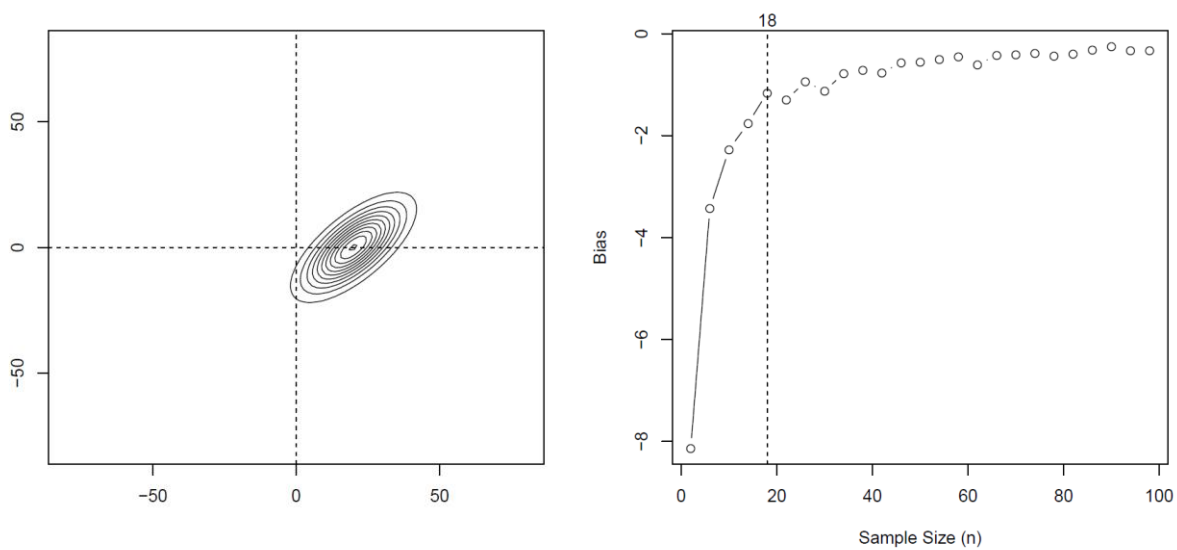


Figure 56 Simulation 24 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

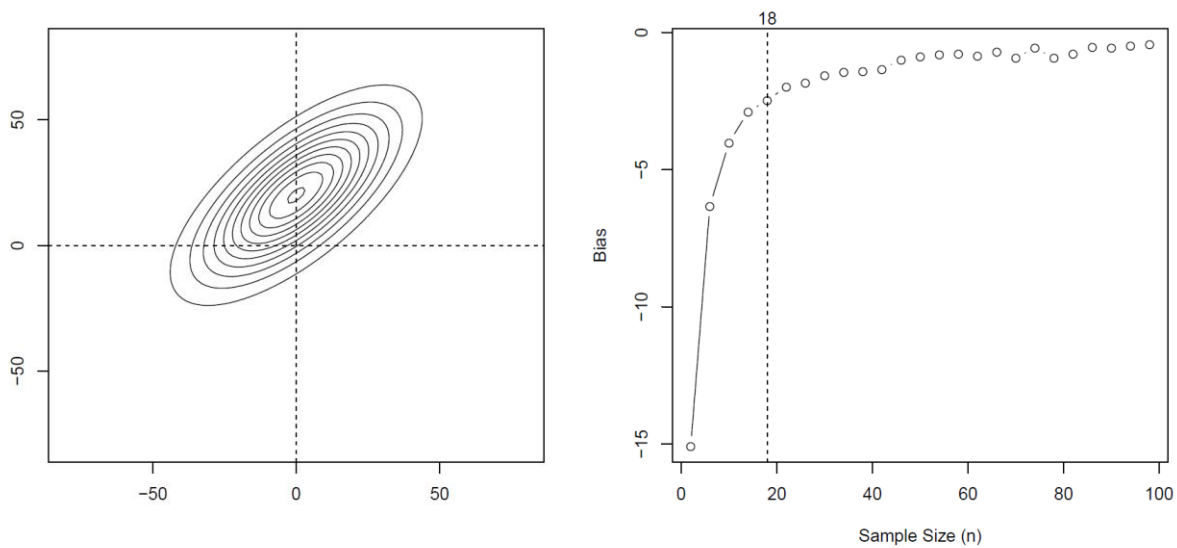


Figure 57 Simulation 25 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

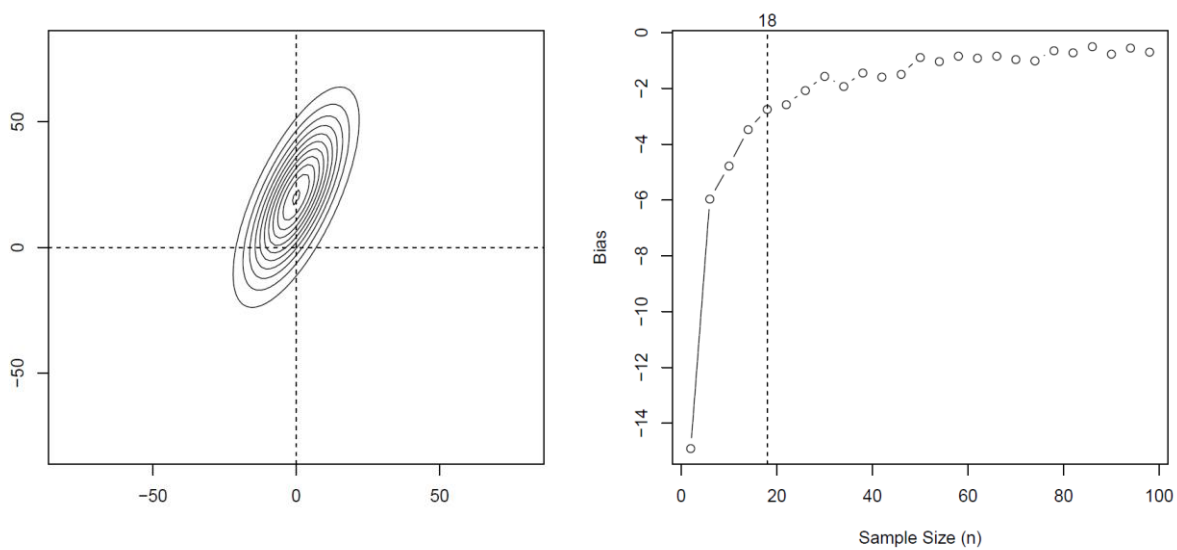


Figure 58 Simulation 26 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

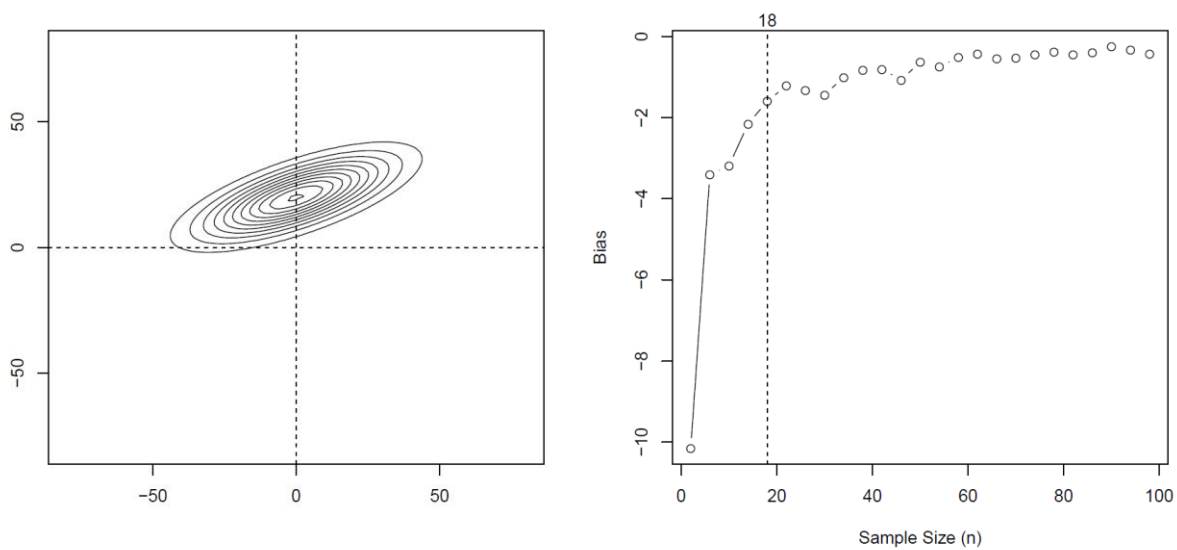


Figure 59 Simulation 27 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

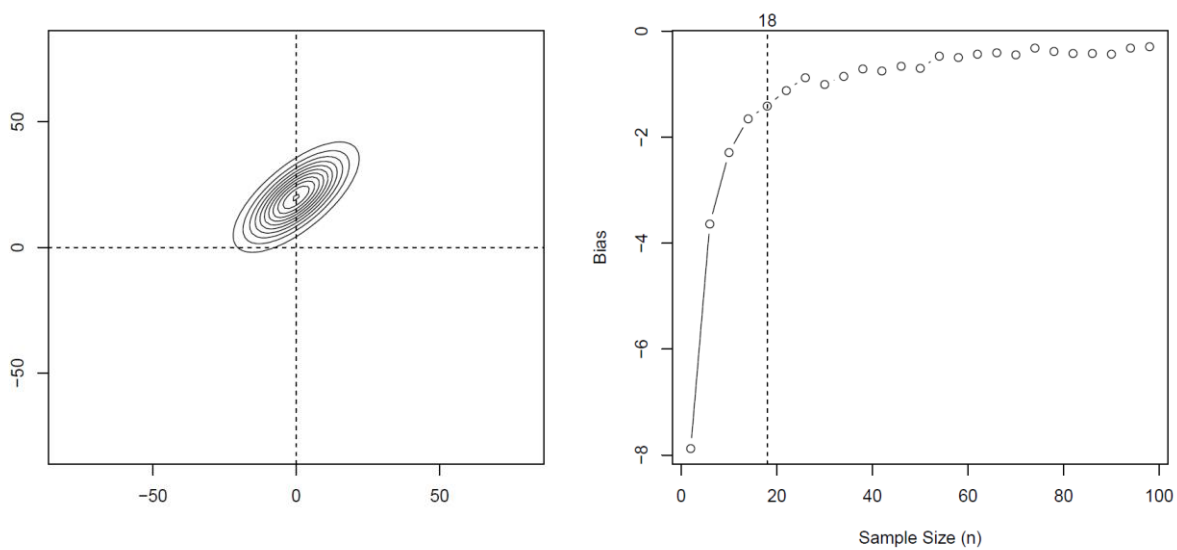


Figure 60 Simulation 28 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

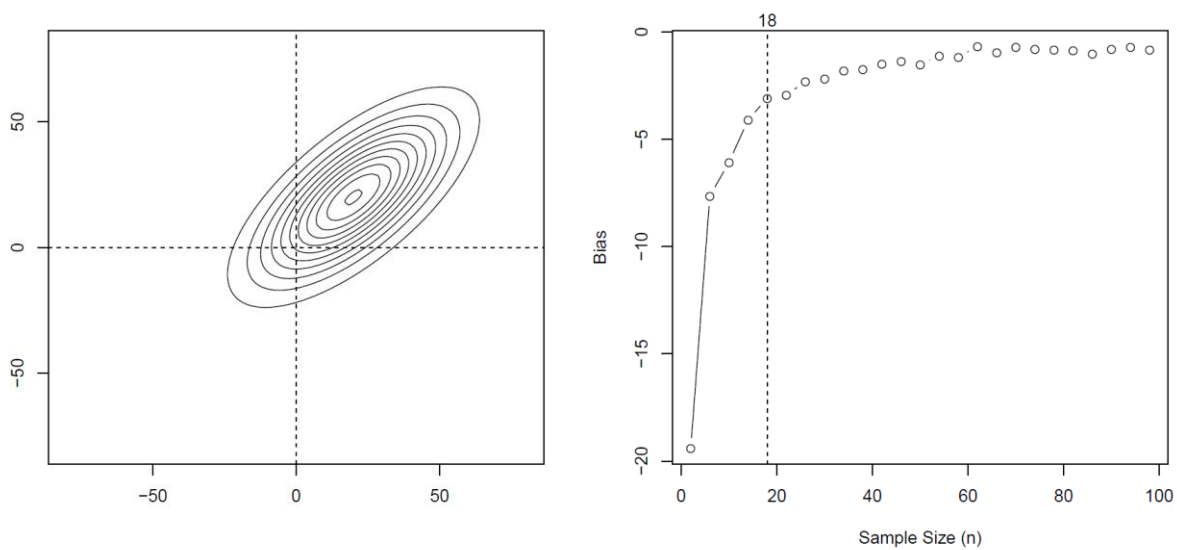


Figure 61 Simulation 29 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

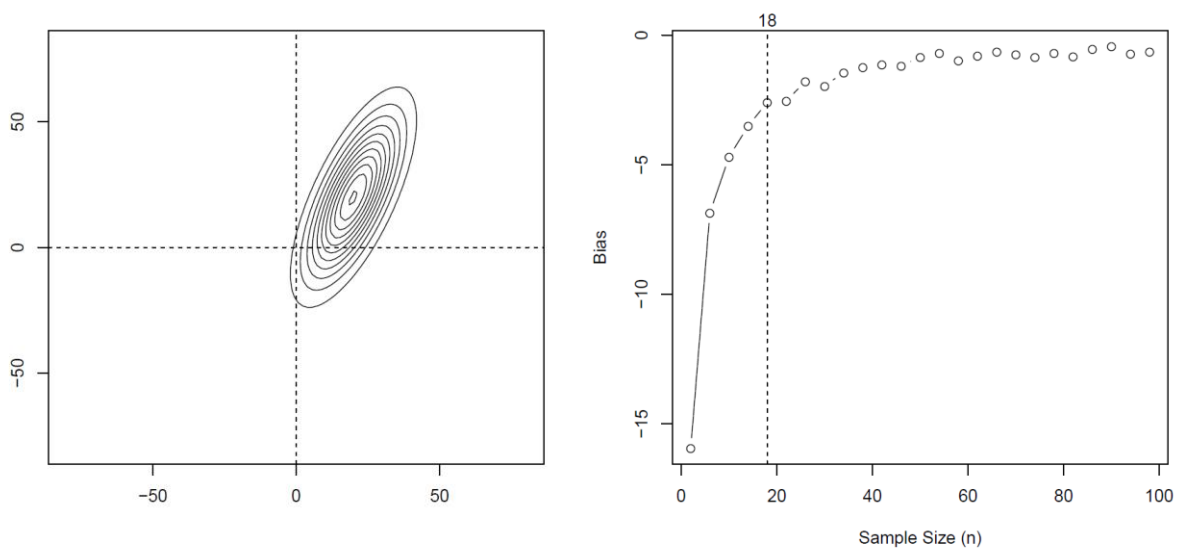


Figure 62 Simulation 30 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

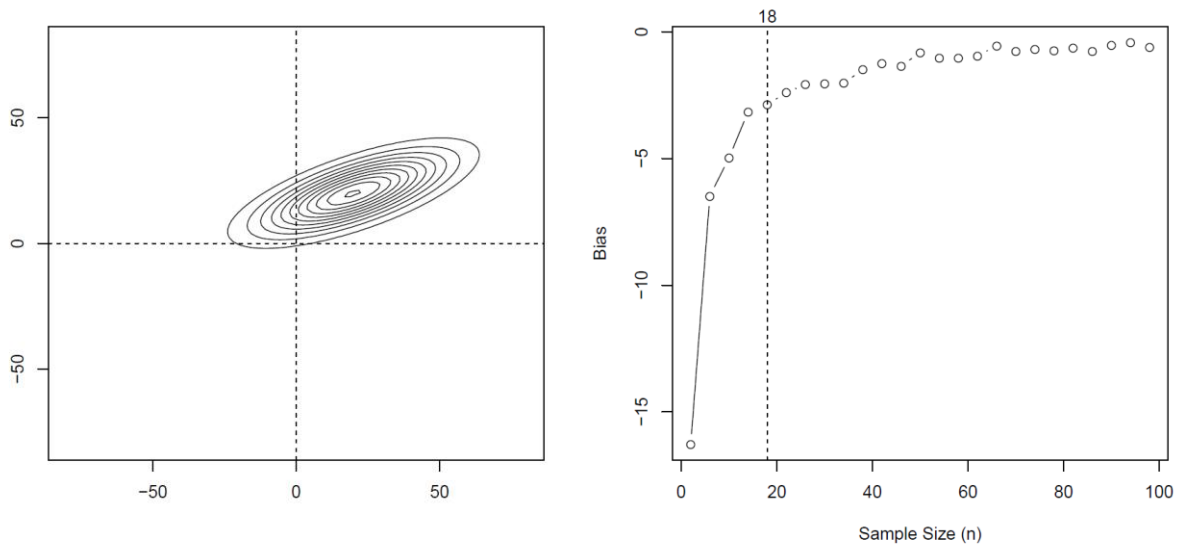


Figure 63 Simulation 31 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

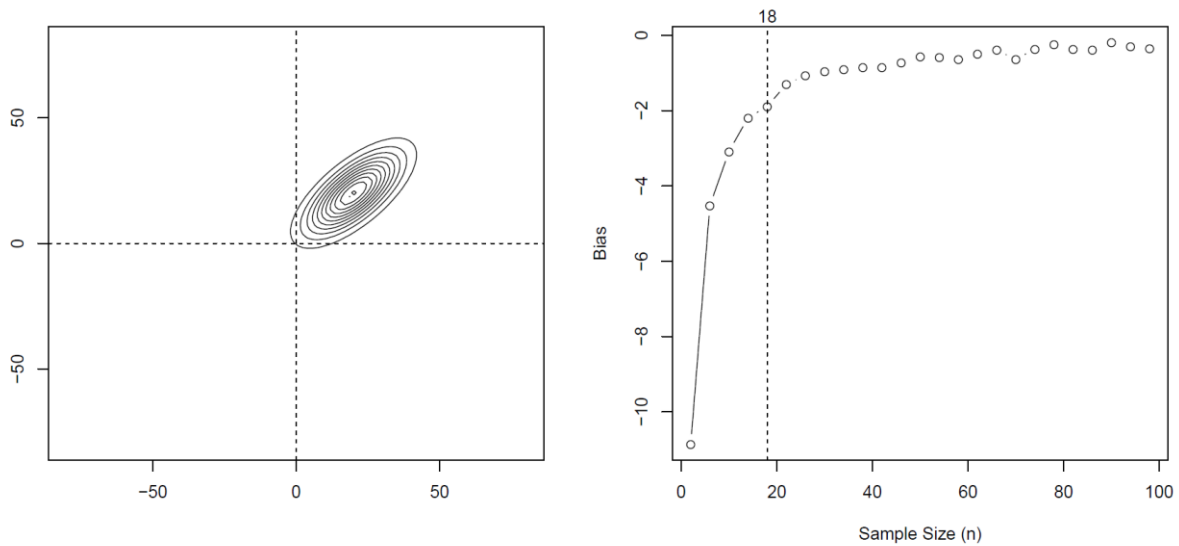


Figure 64 Simulation 32 visual TLE behavior (left), CE90 bias behavior vs. sample size (right).

As would be expected, the bias is negative for CE90 indicating that the estimator is on the low side underestimating the target value of the 90th percentile. The suggested range for sample size on CE90 is about 18 to 30. It seems that a sample of size 18 is sufficient to reach a level of diminishing returns that produces on average a rather small bias from the truth. In one scenario both the x and y standard deviations are 10 feet, this average bias of 2 feet is a 5th of one standard deviation. In other scenarios, the standard deviation for both is 20 feet making this average bias about a 10th of one standard

deviation. By adding more samples there is not much to be gained in reducing the bias, hence in this case it seems sufficient to use 18 as the sample size. Note that skewed distributions are not considered.